numero 228
agosto 1987
anno xx
volume xxxix

La siccità in Africa

Le autorità politiche dei paesi subsahariani dovrebbero tener conto della ricorrenza di questo disastroso evento climatico nei loro piani di sviluppo economico, in modo da stabilizzare la produzione agricola della regione

di Michael H. Glantz

Nelle fotografie riprese nell'Africa subsahariana si vedono spesso bambini denutriti al limite della sopravvivenza, animali scheletrici, rifugiati accalcati nei campi profughi, pozzi asciutti. Queste scene sconvolgenti rappresentano il prezzo della periodica siccità che negli ultimi 20 anni ha attirato sulla regione l'attenzione di tutto il mondo, dando vita a grandi iniziative umanitarie e assistenziali.

Durante ogni episodio di siccità, Governi, agenzie internazionali di assistenza e organizzazioni di carità mettono a punto programmi di aiuto immediato e molto si scrive sull'uso improprio del territorio e sulla desertificazione. Tuttavia, come accadde lo scorso anno, al ritorno delle piogge vi è la tendenza a considerare la siccità un fatto passato e a distogliere l'attenzione dal problema. In realtà, le caratteristiche del clima della regione sono tali da far prevedere che episodi di siccità continueranno a manifestarsi periodicamente. La realizzazione di programmi per ottenere una produzione agricola adeguata nella maggior parte dei paesi della regione subsahariana è strettamente connessa al problema della siccità, che non dovrebbe essere più ignorato nella pianificazione per lo sviluppo. Una chiara conoscenza delle cause della siccità è il primo passo per acquisire una visione globale del problema; in un secondo momento se ne dovrebbero considerare gli effetti e come possano essere ridotti.

È probabilmente inutile cercare un'unica causa per la siccità in Africa. I regimi climatici locali e regionali sono molto diversi perché dipendono da vari processi atmosferici e da caratteristiche topografiche differenti. Nella regione interessata sono presenti, inoltre, società che fanno un diverso uso della terra e ciascun tipo di uso richiede risorse idriche in quantità diversa.

Su scala globale, gli episodi di siccità sono un fenomeno abbastanza comune. Il loro andamento varia considerevolmente di anno in anno; un'annata con relativamente pochi episodi di siccità può essere seguita da un'annata con episodi frequenti. Su scala regionale occorre notare che alcune zone presentano un'unica stagione delle piogge, altre due, che possono cadere sia d'inverno, sia d'estate. Per esempio, il Sahel, nell'Africa occidentale (la zona di transizione tra il Sahara e la savana umida più a sud), ha una stagione secca di otto mesi e una stagione umida di quattro in coincidenza con l'estate dell'emisfero boreale. A livello locale, durante la stagione delle piogge, la variabilità delle precipitazioni, sia nel tempo sia nello spazio, può essere piuttosto elevata.

Una delle maggiori difficoltà che si presentano nell'affrontare il fenomeno della siccità è il suo carattere poco definito. È spesso difficile riconoscere l'inizio e la fine di una fase siccitosa, dal momento che non si presenta molto diversa da normali brevi periodi di aridità. «Il primo giorno senza piogge, in un periodo di bel tempo, contribuisce alla siccità quanto l'ultimo - osservò una volta Ivan R. Tannehill dell'US Weather Bureau - ma nessuno conoscerà con esattezza l'entità del fenomeno fino a che non sia trascorso l'ultimo giorno senza precipitazioni e le piogge siano infine ritornate.»

Le varie popolazioni attribuiscono significati diversi al termine siccità, in funzione dell'importanza e dell'interesse che le precipitazioni rivestono presso di loro. Generalmente la siccità è considerata un fenomeno meteorologico, ma esistono anche la siccità agricola e quella idrogeologica.

La siccità meteorologica può essere valutata in termini del grado di aridità (definito come riduzione percentuale rispetto alla media annuale o stagionale delle precipitazioni calcolata su un lungo

Le dimensioni del lago Ciad, che è suddiviso politicamente tra Camerun, Ciad, Niger e Nigeria, sono diminuite considerevolmente a partire dagli anni sessanta, a causa della prolungata siccità che in Africa ha colpito la regione del Sahel occidentale. Quando la fotografia in alto fu ripresa da un satellite Landsat nel 1972, la superficie del lago era di circa 25 000 chilometri quadrati. Nella fotografia in basso, ripresa da un altro satellite Landsat nel 1979, dopo diversi anni di precipitazioni al di sotto della media, il lago si era ridotto a meno di 2000 chilometri quadrati. Nella seconda immagine sono visibili antiche dune, che sono state a lungo coperte dalle acque.

periodo) e della durata di questa in una determinata regione. Gli intervalli di valori entro cui definiamo la siccità meteorologica sono variabili, dal momento che il suo significato è spesso specifico di una regione e dipende dalle attività umane alle quali vengono riferite le precipitazioni. Una siccità meteorologica è a volte difficile da definire con un buon grado di attendibilità, in parte per la natura del fenomeno e in parte perché in molti paesi africani i dati meteorologici e climatici sono disponibili solo da pochi anni o sono scarsamente affidabili. Inoltre, i soli dati sulle precipitazioni spesso non sono di utilità immediata e diretta per i responsabili politici e per coloro che si occupano di pianificazione agricola, poiché altre variabili possono influire sull'efficacia della pioggia caduta, e fra queste si possono annoverare l'umidità del suolo, la temperatura dell'aria e il tasso di evaporazione.

Si ha siccità agricola quando non vi è un sufficiente livello di umidità nel periodo della crescita e della maturazione delle colture. La distribuzione nel tempo delle piogge durante la stagione agricola è importante quanto il valore assoluto delle precipitazioni mensili o stagionali, dal momento che le colture hanno necessità variabili d'acqua durante il ciclo colturale. M. D. Dennett dell'Università di Reading e i suoi colleghi Jeremy Elston e J. A. Rogers hanno recentemente dimostrato che la distribuzione stagionale delle precipitazioni nel Sahel è cambiata soprattutto per la riduzione delle piogge in agosto, che dovrebbe essere, in media, il mese più piovoso. Soltanto una valutazione a posteriori potrà stabilire se questo fatto sia indice di una tendenza duratura.

Si ha siccità idrogeologica quando la portata dei corsi d'acqua scende al di sotto di una determinata soglia per un certo periodo di tempo. In genere, questo tipo di siccità viene definito come il livello al quale la riduzione della portata è tale da ostacolare in modo significativo alcune attività umane, quali per esempio l'irrigazione delle aree coltivate e la produzione di energia idroelettrica. Dalla fine degli anni sessanta, la portata dei fiumi Niger, Chari e Senegal, in Africa occidentale, è drasticamente diminuita.

Per quanto riguarda la siccità meteorologica, l'Africa occidentale ha conosciuto ultimamente un periodo secco durato 17 anni. Le registrazioni storiche dimostrano come questo sia il terzo episodio di siccità che ha colpito la regione nel XX secolo. Una visione più generale rivela come le variazioni climatiche avvengano a diverse scale temporali, di cui le più rilevanti a questo proposito sono le scale dei millenni, dei decenni e degli anni.

Riferendosi alle variazioni di più lungo periodo, molti ricercatori, allo scopo di spiegare il progressivo instaurarsi di condizioni di aridità nell'Africa subsahariana, si rifanno alla cosiddetta teoria di Milankovich. Nel 1930, l'astronomo serbo Milutin Milankovich ipotizzò che variazioni dell'orbita ellittica della Terra intorno al Sole potessero influenzare il clima. Tali variazioni, che hanno periodi di migliaia di anni, sono causate dall'attrazione gravitazionale esercitata sulla Terra dai pianeti maggiori (si veda l'articolo *L'orbita terrestre e le epoche glaciali* di Curt Covey in «Le Scienze» n.188, aprile 1984).

In conseguenza di ciò, circa 10 000 anni fa l'emisfero boreale riceveva, rispetto a oggi, l'8 per cento in più di radiazione solare d'estate, e l'8 per cento in meno d'inverno, cosicché le estati erano in genere più calde e gli inverni più freddi di quanto non siano ora. La maggiore escursione termica stagionale comportava una intensificazione della circolazione monsonica sia estiva sia invernale nell'emisfero boreale. Da questa circolazione nei due periodi dipende la stagione delle piogge nelle regioni subtropicali.

John E. Kutzbach dell'Università del Wisconsin a Madison e Alayne Street-Perrott dell'Università di Oxford hanno elaborato un modello climatico teorico, che, tenendo conto fra l'altro di queste variazioni, è in grado di riprodurre abbastanza bene le variazioni di precipitazioni nelle regioni subtropicali dell'Africa settentrionale, dell'Asia meridionale e sudorientale e dell'America centrale, che sono testimoniate dall'abbassamento del livello dei laghi negli ultimi 18 000 anni. I livelli dei laghi erano generalmente più alti nel periodo compreso fra 10 000 e 5000 anni fa; l'abbassamento che si riscontra nella maggior parte dei casi a partire da allora evidenzia uno spostamento molto graduale verso un clima più secco. Oggi la Terra raggiunge il perielio durante l'inverno boreale, ma nei prossimi millenni il momento del suo massimo avvicinamento al Sole cadrà nuovamente nel corso dell'estate boreale. Questo fatto dovrebbe rendere più intensa l'attività monsonica (e di conseguenza le piogge) nelle regioni tropicali.

Se si considerano periodi di decenni o di anni, la ricerca delle cause della siccità in Africa si incentra su fattori sia naturali sia antropici. I fattori naturali comprendono le fluttuazioni climatiche casuali di breve periodo, le oscillazioni di lungo periodo, le variazioni di temperatura delle acque superficiali dell'oceano Atlantico, gli effetti del fenomeno di El Niño-Southern Oscillation (ENSO) e le condizioni climatiche anomale a esso collegate, le quali rappresentano un esempio di ciò che i meteorologi chiamano «teleconnessioni», ovvero correlazioni a grande distanza (si veda l'articolo *El Niño* di Colin S. Ramage in «Le Scienze» n.216, agosto 1986). I fattori antropici che si possono tenere in considerazione comprendono l'aumento nell'atmosfera di gas che inducono l'effetto serra, come l'anidride carbonica e altri gas in tracce prodotti da attività umane, e gli interventi che modificano la superficie terrestre.

La variabilità del clima africano è espressa come scarto medio annuo dai valori normali di piovosità. In quelle zone del continente in cui prevalgono regimi pluviometrici irregolari e il valore medio della piovosità è basso, gli episodi di siccità costituiscono una caratteristica permanente del clima. Questo accade in particolare nelle aree continentali in colore rosso chiaro.

Le suddivisioni geografico-climatiche dell'Africa occidentale, a partire dall'area indicata in rosso chiaro, sono il Sahara, la subregione saharo-saheliana, il Sahel, la subregione sudano-saheliana e la regione sudanese. La media annuale delle precipitazioni è indicata in millimetri.

Nella variabilità climatica di breve termine, le fasi siccitose in regioni aride e semiaride possono essere considerate «normali». In queste zone la semplice valutazione statistica della media annuale delle precipitazioni è insufficiente, poiché un numero limitato di anni con precipitazioni abbondanti è mediato da un maggior numero di anni con scarse precipitazioni. In realtà, non vi sono molti anni in cui le precipitazioni sono vicine al valore medio. La valutazione della siccità in queste regioni basata semplicemente sulle differenze con le medie annuali delle precipitazioni porta, di conseguenza, a conclusioni sbagliate. Per descrivere adeguatamente le caratteristiche del regime pluviometrico di una particolare regione africana, si devono prendere in considerazione altri parametri statistici, quali la mediana delle precipitazioni (il valore che si trova al centro di una serie ordinata di quantità pluviometriche), l'intervallo (il valore più alto e quello più basso) e il valore modale (quello che ricorre con maggior frequenza).

Gli studi paleoambientali dimostrano che, in varie parti dell'Africa subsahariana, si sono succeduti nel corso di migliaia di anni lunghi periodi di clima umido e altri di clima secco. Molti ricercatori hanno tentato di utilizzare dati statistici e storici per individuare cicli di siccità in alcune regioni, ma i loro risultati non hanno ricevuto finora una conferma scientifica; al contrario, hanno suggerito che l'insorgere di fasi secche possa essere un fenomeno non periodico, anche se ricorrente.

Per quanto riguarda le variazioni climatiche di lungo periodo, all'inizio degli anni settanta i climatologi si chiedevano se le temperature medie su tutta la Terra fossero in aumento o in diminuzione. I sostenitori dell'ipotesi del raffreddamento prevedevano l'avvicinarsi di una nuova glaciazione, dal momento che l'attuale periodo interglaciale ha avuto una durata approssimativamente uguale a quella dei periodi precedenti (compresa fra 10 000 e 15 000 anni). Inoltre, si sottolineava che, nell'ultimo mezzo milione di anni, temperature elevate come quelle che hanno caratterizzato il XX secolo hanno predominato per soli 25 000 anni circa.

L'ipotesi del riscaldamento ha riscosso invece i maggiori favori nello scorso decennio. Osservazioni dettagliate hanno rivelato che la tendenza all'abbassamento della temperatura, iniziata intorno al 1940 nell'emisfero boreale, si è invertita a metà degli anni settanta. La tendenza di più lungo periodo verso l'innalzamento delle temperature globali, incominciata nei primi anni del secolo, è stata attribuita principalmente all'aumento dell'anidride carbonica nell'atmosfera. Non è chiaro se un aumento generale della temperatura su scala mondiale possa portare a una maggiore piovosità in quelle regioni africane attualmente considerate esposte alla siccità.

El Niño è la temporanea intrusione di acque calde superficiali nel Pacifico orientale, nella zona equatoriale al largo delle coste del Perù e dell'Ecuador. È una manifestazione locale del fenomeno chiamato Southern Oscillation (oscillazione australe), che consiste in una fluttuazione correlata dei valori della pressione atmosferica a livello del mare fra il Pacifico occidentale e quello orientale, a latitudini equatoriali. I fenomeni ENSO sono stati associati a siccità e ad altre anomalie climatiche su scala mondiale. L'evento del 1982-1983 è stato il più intenso da almeno un secolo, per il notevole aumento di temperatura delle acque superficiali, l'estensione geografica e l'impatto sociale. Eugene M. Rasmusson dell'Università del Maryland a College Park sostiene che esista una stretta correlazione tra gli eventi ENSO e le precipitazioni nell'Africa sudorientale (Mozambico e Zimbabwe), dove, su 28 eventi ENSO, 22 hanno coinciso con una riduzione delle precipitazioni. Le correlazioni tra i fenomeni ENSO e le precipitazioni in Etiopia, nel Sahel e in Africa orientale

(Tanzania, Kenya e Uganda) sembrano essere molto più deboli.

I periodi di siccità nell'Africa occidentale possono essere meglio spiegati considerando le variazioni di temperatura delle acque superficiali dell'oceano Atlantico. I modelli climatici elaborati da D. E. Parker, C. K. Folland e T. N. Palmer del British Meteorological Office suggeriscono che, secondo le parole degli autori, «acque più calde del normale nella zona tropicale meridionale dell'Atlantico, in particolare nel Golfo di Guinea, tendono a favorire l'instaurarsi di condizioni di aridità nel Sahel durante la stagione delle piogge, come conseguenza di variazioni della circolazione atmosferica e del trasporto di umidità ai tropici».

La principale attività umana che può destare preoccupazioni è l'uso di combustibili fossili, giunto oggi a livelli senza precedenti nella storia. Un numero sempre maggiore di scienziati sostiene che l'aumento della concentrazione atmosferica di anidride carbonica e di altri gas che favoriscono l'effetto serra, quali metano, ozono, fluorocarburi e ossidi di azoto, riscaldi gli strati più bassi dell'atmosfera. Questi gas sono trasparenti alle radiazioni luminose visibili, ma assorbono o rimandano verso la Terra le radiazioni infrarosse, di lunghezza d'onda più elevata, riflesse verso l'atmosfera dalla superficie terrestre.

Un riscaldamento degli strati più bassi dell'atmosfera potrebbe influenzare i processi idrogeologici e la distribuzione spaziale delle precipitazioni, sebbene gli effetti su scala regionale del fenomeno non siano stati ancora ben interpretati. Nonostante ciò, alcuni esperti nel campo dei processi atmosferici hanno avanzato l'ipotesi che la recente lunga fase secca che ha interessato l'Africa possa essere una prima manifestazione a livello regionale delle conseguenze del riscaldamento dell'atmosfera.

Le modifiche apportate alla superficie terrestre con interventi di deforestazione, pascolo intensivo e desertificazione, nonché l'abbattimento di alberi per ottenere legname da costruzione o da utilizzare come combustibile rappresentano ulteriori attività umane che potrebbero essere responsabili di episodi di siccità. Queste alterazioni possono aumentare l'albedo della superficie terrestre e far sì che questa, assorbendo meno radiazione solare, diventi più fredda. Questo effetto comporta a sua volta modifiche degli strati più bassi dell'atmosfera. La presenza di aria fredda presso la superficie terrestre, a cui si sovrappone uno strato di aria calda, riduce i processi convettivi dell'atmosfera, inibendo la formazione di nubi e di precipitazioni.

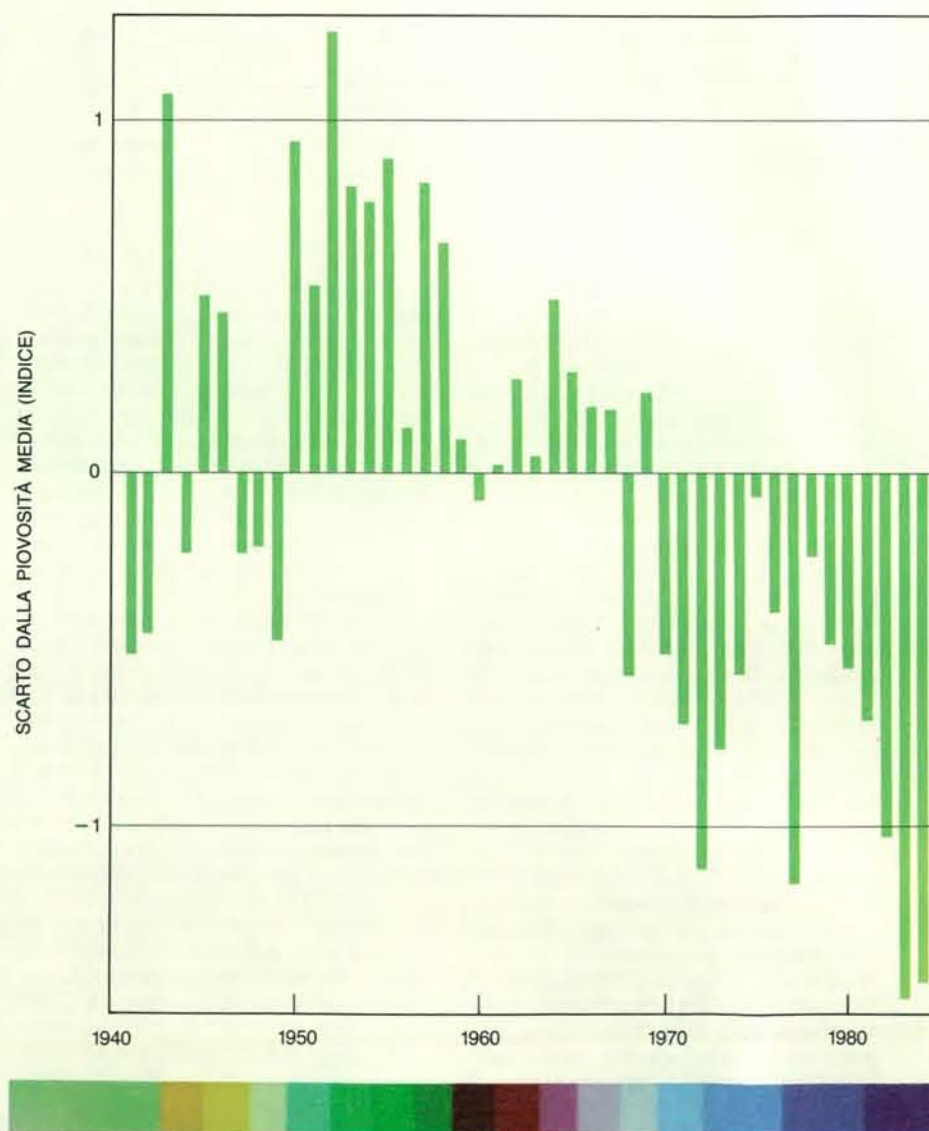
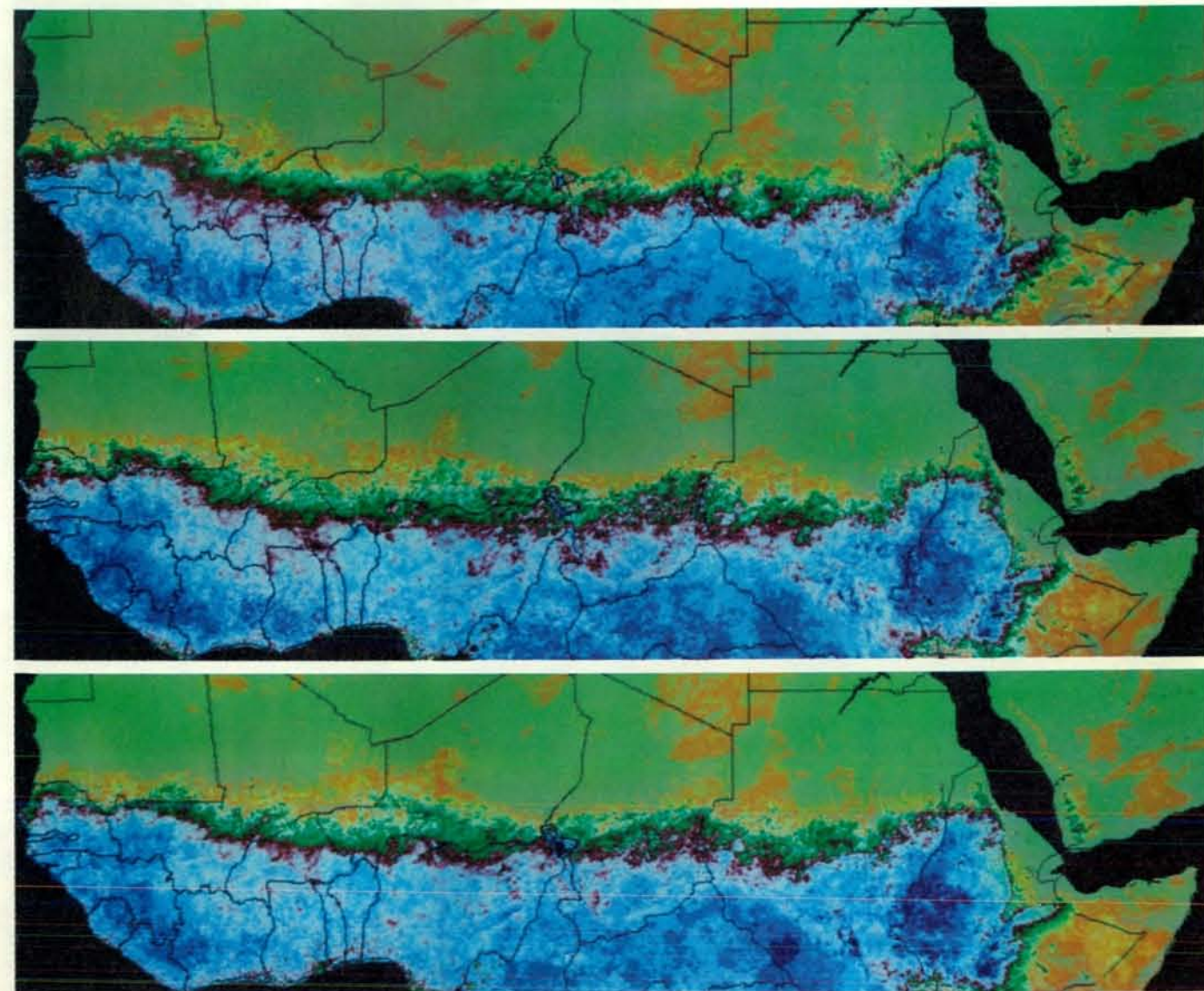
Una decina di anni fa, Jule G. Charney del Massachusetts Institute of Technology suggeriva, insieme ad altri, che l'aumento dell'albedo della superficie terrestre potesse aggravare la siccità su scala regionale. Secondo Charney, i periodi secchi diventano più intensi e ripetitivi perché un numero crescente di persone è costretto a trovare sostentamento in una base di risorse alimentari sempre più ridotta, e risulta inevitabile un'ulteriore degradazione della copertura vegetale. Tuttavia, tale ipotesi è stata di recente messa in discussione. Ricerche storiche sui cambiamenti ecologici nel Sahel occidentale hanno rivelato che le variazioni di albedo hanno avuto una rilevanza inferiore rispetto a quanto fosse stato previsto nei modelli elaborati al calcolatore. Queste scoperte dimostrano che, anche quando si manifesta un aumento dell'albedo, questo fatto probabilmente non è in grado di comportare effetti pronunciati a livello regionale.

Si è anche ipotizzato che le modificazioni subite dalla superficie terrestre possano influire sulle precipitazioni in altri modi. Per esempio, potrebbero ridurre il numero dei nuclei ghiacciati di condensazione derivanti dalla decomposizione di foglie e di altri residui vegetali. Tali nuclei favoriscono la condensazione del vapore acqueo in pioggia. È stata avanzata l'ipotesi che i nuclei organici siano più attivi nel processo di condensazione di quanto non lo siano nuclei inorganici come il pulviscolo, dal momento che questi ultimi raggiungono il punto di congelamento a una temperatura delle nubi molto più bassa.

La desertificazione è un'altra modificazione della superficie terrestre che potrebbe influire sul regime pluviometrico, in quanto può determinare un aumento della quantità di pulviscolo negli strati più bassi dell'atmosfera. Il pulviscolo assorbe e disperde la radiazione solare, cosicché la parte superiore degli strati ricchi di pulviscolo si riscalda, impedendo a una parte della radiazione di raggiungere la superficie terrestre, relativamente più fredda. Anche in questo caso, l'intensità dei fenomeni atmosferici che favoriscono le precipitazioni viene ridotta.

Gli studi delle condizioni geofisiche considerano una sola categoria di possibili cause delle prolungate siccità e delle carestie che hanno colpito l'Africa, ma trascurano una componente importante del problema: le complesse interazioni esistenti tra variazioni climatiche e attività umane. Occorre tener conto di queste interazioni per capire come la siccità incida sull'agricoltura, gli ecosistemi e l'economia.

Nell'Africa subsahariana, dove più dell'80 per cento della popolazione è occupato nell'agricoltura e quindi il suo benessere è direttamente collegato alle precipitazioni, le conseguenze della siccità sono molteplici. Alcuni effetti sono evidentissimi: abbeveratoi completamente asciutti, colture avvizzite e riduzione del foraggio per il bestiame, ma molti effetti, meno ovvi, sono forse altrettanto importanti; fra questi, l'aumento dei prezzi, la crescita delle importazioni di derrate alimentari, il peggioramento della dieta delle popolazioni esposte, le ondate di migrazione dalla campagna verso i centri urbani.

Le migrazioni evidenziano come gli scarsi raccolti e il rapido aumento del prezzo dei cereali sul mercato, dovuti a una prolungata siccità, abbiano conseguenze rovinose sulle economie di villaggio. I contadini più poveri, fortemente indebitati e con scorte di cereali spesso scarse, sono i più duramente colpiti. Quando la siccità agricola continua per alcune stagioni, la popolazione tende ad abbandonare i villaggi. Gli uomini sono i primi ad andarsene, alla ricerca di un lavoro che permetta loro di guadagnare abbastanza per comprare del cibo; donne e bambini li seguono successivamente, per ricomporre la famiglia o spinti semplicemente dalla necessità di cibo. Se la siccità continua per un lungo periodo, come è accaduto in alcune regioni dell'Africa subsahariana, gli emigranti finiscono stremati nei campi profughi, totalmente dipendenti dagli aiuti alimentari. Molti di questi spostamenti forzati diventano permanenti; la popolazione occupata nell'agricoltura diminuisce progressivamente e, di conseguenza, la produzione di derrate alimentari pro capite subisce un rapido declino.

Sulla base dei dati raccolti da 20 stazioni pluviometriche, Peter J. Lamb, dell'Illinois State Water Survey, ha costruito un indice delle precipitazioni di una regione dell'Africa occidentale comprendente il Sahel. È evidente che le precipitazioni dal 1968 al 1985 sono rimaste sotto la media. La situazione è in parte migliorata lo scorso anno e gli effetti sulla vegetazione si osservano nei dati da satellite elaborati dalla Food and Agriculture Organization (FAO) e dalla National Aeronautics and Space Administration (NASA). Dall'alto in basso, le fotografie illustrano la situazione in agosto e settembre del 1984, 1985 e 1986. Il colore verde chiaro indica la minore copertura vegetale, il blu scuro la maggiore.

Nell'Africa subsahariana la produzione di derrate alimentari pro capite è diminuita a partire dall'inizio degli anni sessanta, in parte come conseguenza delle scarse precipitazioni. La produzione dell'Africa subsahariana (*in rosso*), con esclusione del Sud Africa, viene qui messa a confronto con quella dell'America Latina (*in marrone*) e di sei nazioni dell'Asia meridionale (*in grigio*). I dati si basano su informazioni raccolte ed elaborate dall'US Department of Agriculture.

I pastori spesso sono i primi ad avvertire le conseguenze di un periodo di siccità, poiché di solito vivono ai margini del deserto. Nel Sahel, quando le piogge si arrestano più a sud del normale, i pascoli si disseccano e le mandrie hanno a disposizione meno foraggio. I pastori sono costretti a trovare nuove aree di pascolo per il bestiame. In condizioni di estrema siccità, molti animali muoiono, per mancanza di vegetazione oltre che di acqua, perché il bestiame pascola intensivamente intorno ai punti d'acqua permanenti o semipermanenti e distrugge l'equilibrio tra vegetazione e risorse idriche. Molti pastori finiscono nei campi profughi insieme ai contadini poveri, oppure nei centri urbani dove, abituandosi alla vita cittadina, perdono ogni desiderio di ritornare ai pascoli a lottare contro le avversità del clima.

Nei centri urbani e nei campi profughi, i rifugiati finiscono con il preferire cereali importati, come frumento e riso, ai tradizionali miglio e sorgo. Perciò, sul lungo periodo, un effetto meno evidente della siccità consiste in un aumento della domanda di cereali di importazione, che provoca gravi difficoltà al già precario bilancio degli scambi con l'estero di un paese in via di sviluppo.

Per far fronte a tali costi, le autorità governative devono cercare di aumentare gli scambi con l'estero. Una delle soluzioni più diffuse è la produzione di derrate da destinare all'esportazione, quali cotone, arachidi e caffè. La produzione di queste colture viene ottenuta a spese di quelle destinate all'alimentazione locale, dal momento che nelle aree più fertili e con una maggiore disponibilità di risorse idriche le colture alimentari vengono sostituite con quelle destinate all'esportazione. È significativo che nel Sahel occidentale e in Etiopia, anche durante le gravi siccità dei primi anni settanta e ottanta, la produzione e l'esportazione delle colture commerciali si siano mantenute costanti o siano aumentate mentre la produzione delle colture tradizionali è notevolmente diminuita.

Ogni anno in molte parti del mondo hanno luogo periodi di siccità, che però non provocano necessariamente carestie o grave penuria di generi alimentari. Brasile, India, Indonesia e Kenya sono esempi di paesi che, in anni recenti, hanno fatto fronte in vari modi alle difficili condizioni poste dalla siccità.

Solo pochi dei paesi africani, che hanno accusato scarsità di riserve alimentari nel triennio 1982-1984, hanno sofferto vere e proprie carestie. I paesi che effettivamente sono stati colpiti (Mozambico, Angola, Sudan, Ciad ed Etiopia) dovevano fronteggiare non solo gravi siccità, ma anche conflitti interni. Questo fatto mette in evidenza come la siccità di per sé non determini carestie, ma aggravi drasticamente altri problemi presenti all'interno di una società.

È possibile proporre proiezioni più attendibili dell'impatto sociale della siccità utilizzando dati teorici relativi a situazioni analoghe o analizzando esperienze simili di altre zone. Un paese con conflitti interni ha un'alta probabilità di essere esposto al rischio di carestie al verificarsi di una siccità, come dimostra l'esperienza recente dell'Etiopia. Inoltre, i Governi che destinano i terreni migliori alle colture per l'esportazione costringono i piccoli agricoltori e i pastori a spostarsi alla ricerca di sostentamento in aree relativamente più marginali. Nel momento in cui vengono messi a coltura terreni aridi e semiaridi poco adatti a produzioni agricole che necessitano di piogge costanti, vi saranno maggiori probabilità che la siccità impedisca le coltivazioni, portando alla desertificazione. Questo accadrà non perché siano necessariamente cambiate la quantità di pioggia e la sua distribuzione, ma perché le nuove attività richiedono dotazioni idriche di lungo periodo superiori a ciò che le precipitazioni possono fornire.

Sebbene l'irrigazione sia stata considerata un rimedio promettente dai Governi africani e dagli stranieri che hanno fornito gli impianti, essa si è rivelata spesso inefficace. L'irrigazione costa: le colture che si attuano su terreni irrigui non richiedono soltanto infrastrutture come pompe, tubazioni e canali, ma anche spese supplementari per l'acquisto di concimi, erbicidi e insetticidi. Le colture tradizionali destinate all'alimentazione locale non possono di solito sopportare questi ulteriori costi perché, nella maggior parte dei paesi africani, le politiche statali, mantenendo basso il prezzo dei generi alimentari, riducono anche i profitti degli agricoltori. Tutto ciò fa sì che i Governi sviluppino progetti di irrigazione soltanto per quelle colture, come cotone e zucchero, che permettono di incrementare gli indispensabili scambi con l'estero. Oggi, in Africa, l'irrigazione è diffusa soltanto su una parte limitata della superficie agricola ed è poco probabile che possa venire utilizzata più estesamente nelle tradizionali colture per l'alimentazione locale.

Altri rimedi proposti per combattere la siccità includono vari progetti volti a modificare direttamente il clima e le condizioni meteorologiche. Essi possono essere classificati in base al tipo di intervento previsto. Alcuni possono riguardare la vegetazione (per esempio, la costituzione di fasce di alberi e il rimboschimento delle aree desertificate); alcuni possono modificare la circolazione atmosferica (per esempio, con la creazione di laghi artificiali in antichi bacini per far aumentare l'umidità atmosferica tramite l'evaporazione dell'acqua); altri possono influire sulle precipitazioni (per esempio, con l'inseminazione di fronti monsonici di grande scala e di formazioni nuvolose di scala locale). Tali progetti tecnologici, basati su ipotesi, sia pure scientifiche, sono spesso di valore opinabile e possono servire soltanto a mascherare processi ecologici e sociali più profondi, nonché a generare illusioni.

Le sole scelte plausibili a medio termine sono quelle proposte per alleviare i danni provocati dalla siccità sulle società africane. In questa categoria di opzioni possibili rientra la sostituzione, durante i periodi di siccità, delle colture destinate all'esportazione con colture che servano all'alimentazione della popolazione locale oppure l'assicurazione che la valuta straniera, ricavata con i prodotti esportati, venga impiegata nell'acquisto di alimenti per le popolazioni colpite; inoltre, gli aiuti volontari forniti devono essere pertinenti alle necessità dei beneficiari ed è indispensabile tener conto della situazione agricola e climatica ogni volta che si progettano nuove strutture produttive agricole. In particolare, ogni scelta deve essere basata sul convincimento che la siccità meteorologica non causa di per se stessa i disastri che le si addebitano, ma che, combinandosi con problemi sociali, economici o politici preesistenti, può portare alla distruzione della capacità produttiva agricola di una società.

Se si considera il lungo termine, vi sono altre strade percorribili, come la diffusione, da parte di agenzie quali per esempio la World Bank, di un convincente programma educativo sul problema della siccità, intesa come principale forza destabilizzante e ricorrente, con la quale le autorità governative devono periodicamente misurarsi. È anche importante rendere consapevoli i responsabili politici dei paesi esposti di come la siccità possa diventare un fattore limitante lo sviluppo. Questo programma educativo dovrà essere portato avanti senza interruzioni, dato che spesso i Governi restano in carica per un tempo inferiore al periodo compreso tra due siccità.

Anche per i politici stessi è importante adottare misure che limitino l'espansione delle coltivazioni e del pascolo su terreni non adatti e in aree ai margini delle zone piovose. Quando le coltivazioni e il pascolo si diffondono nelle terre marginali, gli effetti della siccità agricola (tra cui l'avvizzimento delle colture e l'inaridimento dei terreni) generano carenze alimentari più gravi e una desertificazione più estesa.

I Governi, inoltre, potrebbero sfruttare in maniera più ampia nella programmazione agricola i loro servizi meteorologici. Infatti questi servizi possono essere di grande aiuto alle autorità dimostrando come la conoscenza dei dati meteorologici permetta di operare scelte più razionali per lo sviluppo agricolo. Gli stessi servizi possono essere utilizzati anche per poter disporre della segnalazione tempestiva della presenza di condizioni che possono portare a una carestia.

È sempre più chiaro che la sola valutazione degli aspetti fisici della siccità (portata del fenomeno, durata, intensità e area geografica colpita) non è in grado di spiegare fino in fondo perché periodi aridi con caratteristiche apparentemente simili comportino effetti variabili da un paese all'altro e perfino, di volta in volta, nella stessa regione. Una migliore comprensione del fenomeno si avrà soltanto da ricerche multidisciplinari che prendano in considerazione anche fattori sociali, economici e culturali.

I meccanismi molecolari della fotosintesi

È possibile ottenere un quadro molto preciso dello svolgimento di questo processo e dei compiti delle varie molecole utilizzando congiuntamente tecniche spettroscopiche, cristallografia a raggi X e genetica molecolare

di Douglas C. Youvan e Barry L. Marrs

La fotosintesi, il processo mediante il quale la luce solare viene trasformata nell'energia necessaria allo svolgimento delle funzioni vitali degli organismi, è la chiave di volta della vita sulla Terra. L'energia catturata dalla fotosintesi serve non solo a nutrire gli organismi fotosintetizzanti, bensì anche gli animali che si cibano di questi organismi, gli altri animali che a loro volta li divorano e così di seguito, lungo la catena alimentare. Come si svolge la fotosintesi? A livello molecolare, quali sono le interazioni che captano l'energia luminosa del Sole e la trasformano in energia vitale? Qual è la struttura, cioè l'organizzazione spaziale, delle molecole interessate? E come può, quella particolare struttura, originare la velocità e l'efficienza del processo fotosintetico?

Queste domande fondamentali cominciano ora ad avere risposta. Grazie a singoli contributi di ricercatori che operano in campi disparati come la spettroscopia, la cristallografia a raggi X e la genetica molecolare, i meccanismi molecolari di un tipo di fotosintesi (relativa ad alcuni batteri) sono stati descritti con grande ricchezza di particolari.

Ogni campo di ricerca fa luce su un diverso aspetto del processo. Gli spettroscopisti hanno determinato la sequenza e la collocazione temporale degli eventi molecolari che si svolgono nelle «reazioni alla luce» della fotosintesi (il primo stadio del processo) e la velocità con cui procede ogni interazione. I cristallografi hanno visualizzato con i raggi X la struttura spaziale del centro di reazione della fotosintesi (dove si svolgono le reazioni alla luce) e hanno scoperto come le molecole presenti in esso sono allineate l'una rispetto all'altra. Gli esperti di genetica molecolare, d'altra parte, hanno localizzato e analizzato (e possono oggi manipolare) i geni che dirigono la sintesi dei principali componenti del centro di reazione.

La conoscenza delle interazioni molecolari, della struttura e della base genetica del centro di reazione della fotosintesi permette di porre domande più particolareggiate sulla sua funzione. In che modo la presenza di ciascuna molecola contribuisce al funzionamento di tutto il meccanismo? Perché alcuni stadi delle reazioni alla luce sono più veloci o più lenti di altri? Come si modificherebbe il processo se un particolare elemento avesse una forma o una composizione lievemente diversa da quella che ha?

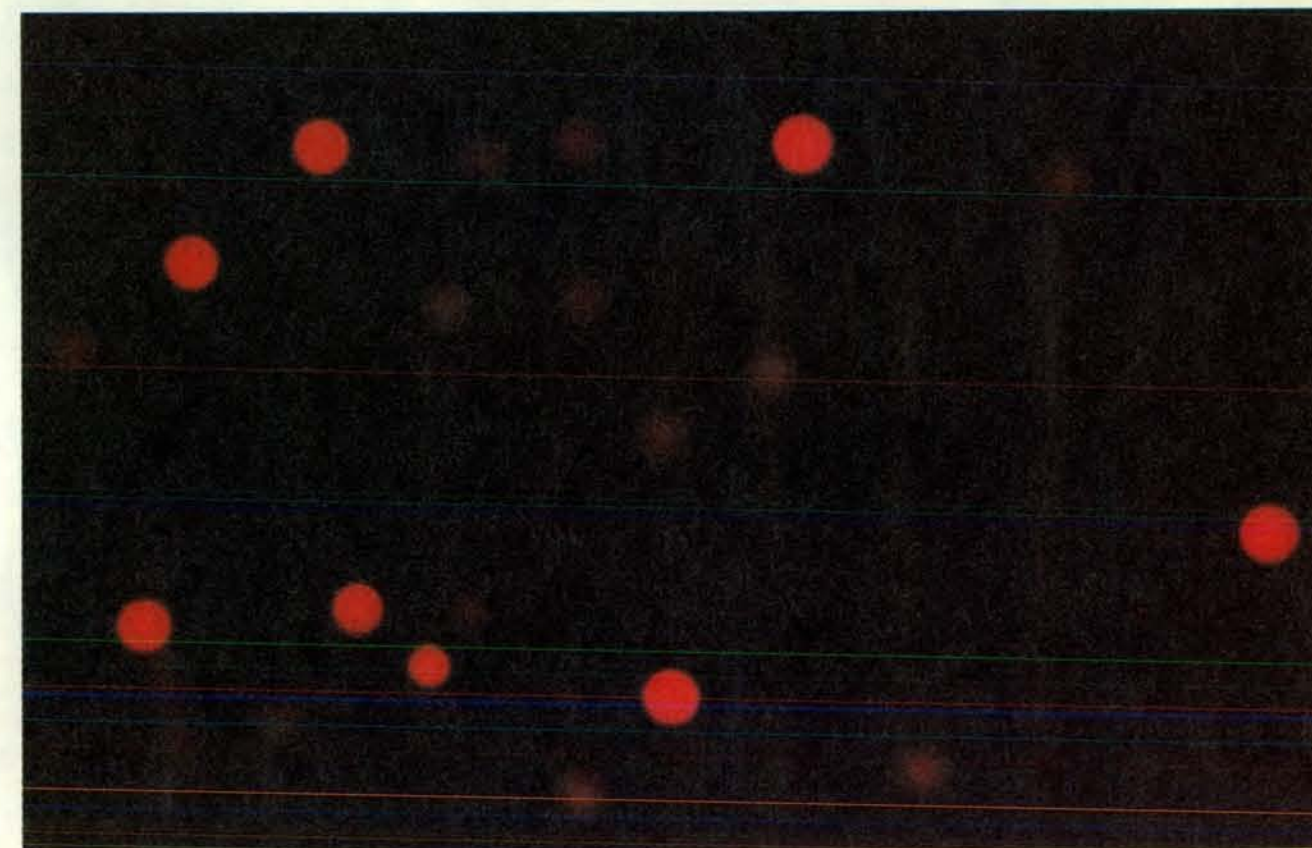
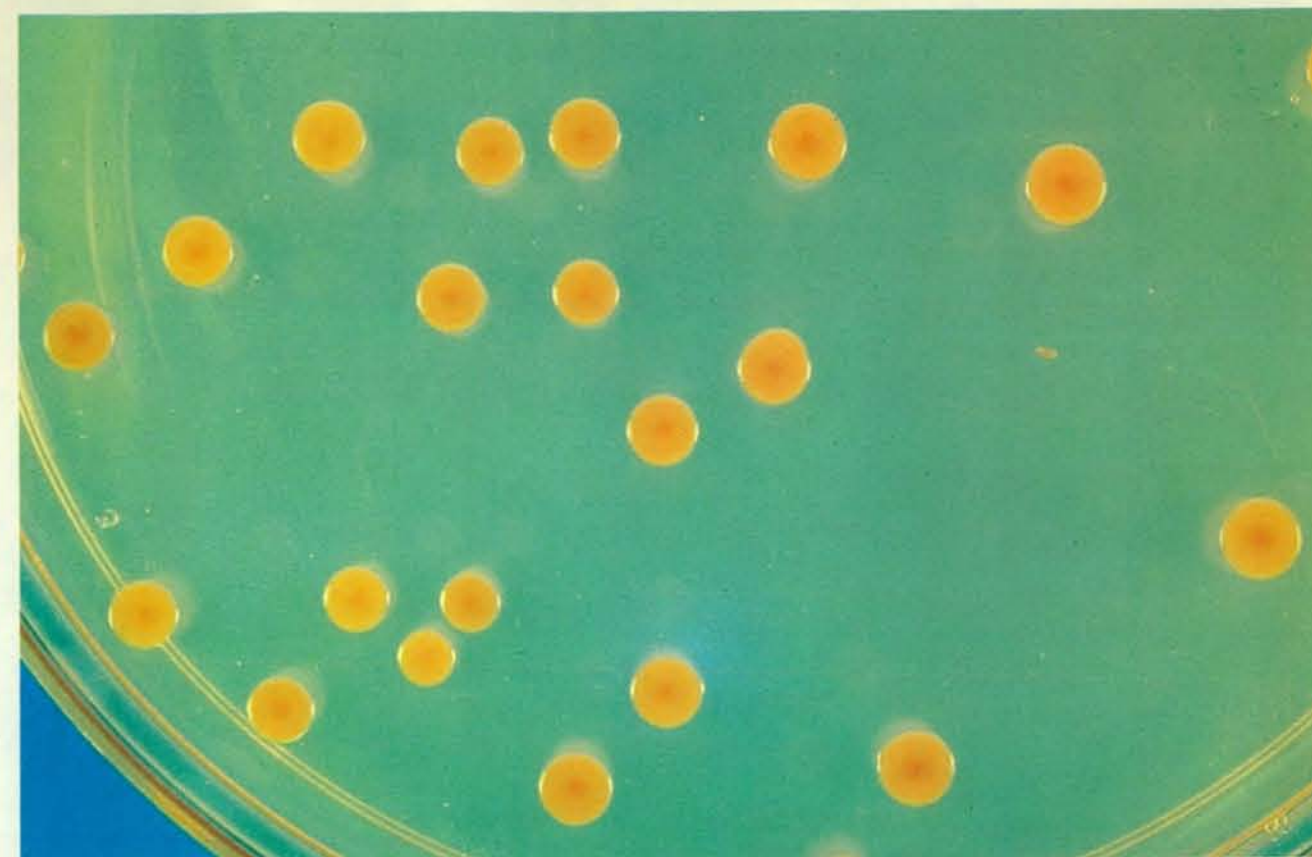
Noi e altri ricercatori ci siamo impegnati per trovare una risposta a queste domande utilizzando i potenti metodi della genetica molecolare. Alterando l'informazione genetica che codifica per certi elementi del centro di reazione, uno di noi (Youvan) ha prodotto batteri fotosintetizzanti con centri di reazione della fotosintesi che differiscono in modi specifici da quelli di batteri che non sono stati manipolati geneticamente. Verificando poi la funzione fotosintetica dei batteri alterati, si vede quale effetto abbia provocato l'alterazione sull'efficienza dell'attività fotosintetica dei batteri.

Queste tecniche permettono ai ricercatori di conoscere i meccanismi della fotosintesi con dovizia di particolari. I risultati, interessanti di per sé, possono anche consentire di progettare e produrre batteri che siano in grado di svolgere la fotosintesi in modi molto specializzati o in ambienti difficili.

Per questi studi vengono utilizzati batteri del genere *Rhodopseudomonas*. Essi svolgono il processo fotosintetico praticamente come i batteri che per primi, più di tre miliardi di anni fa, svilupparono un processo per immagazzinare l'energia solare. La fotosintesi di questi batteri differisce per certi aspetti da quella delle piante superiori; innanzi tutto, non produce ossigeno. Tra i due processi vi sono però molte somiglianze: per esempio, in entrambi si ha la partecipazione di molecole di clorofilla.

Un particolare vantaggio che si ha nello studiare i *Rhodopseudomonas* è che, a differenza delle piante superiori, essi possono ottenere l'energia necessaria per la crescita non solo dalla fotosintesi, ma anche da vari meccanismi indipendenti, potendo così sopravvivere e moltiplicarsi senza fotosintesi. È quindi possibile coltivare colonie di mutanti che presentano difetti nei centri di reazione della fotosintesi e studiare così, in organismi viventi, i modi in cui difetti specifici disabilitano la funzione fotosintetica.

L'interno dei *Rhodopseudomonas* è pieno di vescicole fotosintetiche, che sono sferette cave delimitate da un doppio strato lipidico (il materiale che costituisce le membrane cellulari). I centri di reazione della fotosintesi, costituiti principalmente da proteine, sono inclusi in quel doppio strato: un'estremità del centro di reazione si trova vicino alla superficie esterna del doppio strato e l'altra estremità vicino alla superficie interna.

All'inizio delle reazioni alla luce, un fotone colpisce l'estremità del centro di reazione della fotosintesi vicina alla superficie interna del doppio strato. Un elettrone in questa regione viene eccitato (cioè portato a un livello di energia più alto) e trasporta l'energia del fotone all'altra estremità del centro di reazione (vicina alla superficie esterna del doppio strato) mediante una serie di interazioni chimiche. Un secondo fotone fa sì che un altro elettrone segua lo stesso percorso. Queste interazioni producono una separazione di cariche: gli elettroni, dotati di carica negativa, si spostano vicino alla superficie esterna delle vescicole fotosintetiche, mentre l'eccesso di carica positiva, dovuto alla carenza di elettroni, è localizzato all'estremità delle molecole vicina alla superficie interna. La separazione delle cariche rappresenta un immagazzinamento di energia (infatti,

Una mutazione genetica blocca la funzione fotosintetica in alcune di queste colonie di batteri che appartengono al genere *Rhodopseudomonas*. In condizioni di luce ambiente (*in alto*), i batteri fotosintetizzanti e quelli mutati presentano lo stesso aspetto. Quando, invece, vengono osservati nell'infrarosso vicino (*in basso*), i batteri mutati emettono fluorescenza e appaiono luminosi. Infatti, essi riemettono l'energia luminosa che hanno assorbito, ma che non possono utilizzare. I mutanti sopravvivono perché possono trovare sostentamento in processi diversi dalla fotosintesi. Di conseguenza, è possibile coltivare batteri mutati e studiare gli effetti precisi delle mutazioni nei geni per la fotosintesi.

se elettroni e molecole con carica positiva potessero riunirsi, quell'energia verrebbe liberata). In seguito, essa avvia reazioni chimiche che sono accoppiate a certi processi del metabolismo batterico, ai quali forniscono energia.

Quali interazioni portano gli elettroni da un lato del doppio strato lipidico che delimita le vescicole fotosintetiche all'altro? Hanno risposto a questa domanda gli spettroscopisti, che impiegano le radiazioni elettromagnetiche per sondare la struttura chimica della materia.

Particolarmente utile è risultata la spettroscopia ad assorbimento ottico, nella quale si fa passare un fascio di luce attraverso un campione e si misura quanta di questa luce viene assorbita. Ogni tipo di molecola assorbe solo certe lunghezze d'onda della luce in dipendenza dalla struttura della molecola e dal suo ambiente chimico. Possiamo dire che ogni tipo di molecola ha la propria «impronta digitale» spettroscopica. Lo spettroscopista può, dunque, determinare la composizione di un campione, trovandone lo spettro di assorbimento (che descrive come il campione assorbe radiazioni di differenti lunghezze d'onda).

Quando il centro di reazione della fotosintesi assorbe un fotone, la sua struttura chimica subisce una serie di modificazioni nel corso delle interazioni che trasportano l'elettrone attraverso il doppio strato lipidico. Queste variazioni di tipo chimico possono essere seguite notando come cambia lo spettro di assorbimento del centro di reazione dopo che è stato assorbito un fotone. In effetti, una serie di queste osservazioni permette di risalire alla «via chimica» seguita dall'elettrone per passare da un lato all'altro del doppio strato lipidico.

Molti gruppi di ricercatori hanno compiuto queste osservazioni. Nei loro esperimenti, il fascio sonda (che serve a valutare lo spettro di assorbimento di un campione di centri di reazione della fotosintesi, estratti da una popolazione di batteri) è di solito molto debole per non provocare una grande attività fotosintetica. Viene quindi innescata la fotosintesi nel campione, esponendolo a un lampo intenso di luce laser, di durata inferiore a un picosecondo (10^{-12} secondi). L'impulso laser è così breve che ogni centro di reazione nel campione assorbe al massimo un solo fotone e le reazioni fotosintetiche risultanti sono sincronizzate: in ogni dato momento, ogni centro di reazione sta subendo la stessa trasformazione chimica. Pertanto, lo spettro di assorbimento del campione nel suo insieme riflette lo stato chimico di ogni centro di reazione ed è possibile dedurre, con lampi successivi, il tipo e la collocazione temporale delle reazioni nei primi stadi della fotosintesi.

Un altro metodo spettroscopico è la risonanza di spin elettronico. L'elettrone possiede uno spin e, quindi (come tutte le particelle cariche dotate di spin), dà origine a un campo magnetico; per certi aspetti lo si può considerare un minuscolo magnete. Nella maggior parte dei composti stabili, tutti gli elettroni formano coppie, con i campi magnetici dei due elettroni allineati in versi esattamente opposti. Tuttavia, in alcune molecole vi è un elettrone spaiato. Nella spettroscopia di risonanza di spin elettronico, il ricercatore valuta come l'energia degli elettroni spaiati si modifichi in un campo magnetico che varia rapidamente. I risultati forniscono indizi sull'ambiente molecolare che circonda gli elettroni spaiati.

Negli anni settanta, George Feher e collaboratori dell'Università della California a San Diego e, indipendentemente, James R. Norris e collaboratori dell'Argonne National Laboratory, ricostruirono le fasi iniziali della fotosintesi mediante spettroscopia di risonanza di spin elettronico. Essi poterono osservare, per la prima volta, un segnale proveniente da un elettrone spaiato sulle molecole di clorofilla che avevano assorbito un fotone. L'elettrone era spaiato perché le molecole di clorofilla, assorbendo un fotone, avevano liberato un elettrone. La natura del segnale di risonanza di spin elettronico indicava che l'elemento che aveva assorbito il fotone nel centro di reazione era costituito da due molecole di clorofilla strettamente associate.

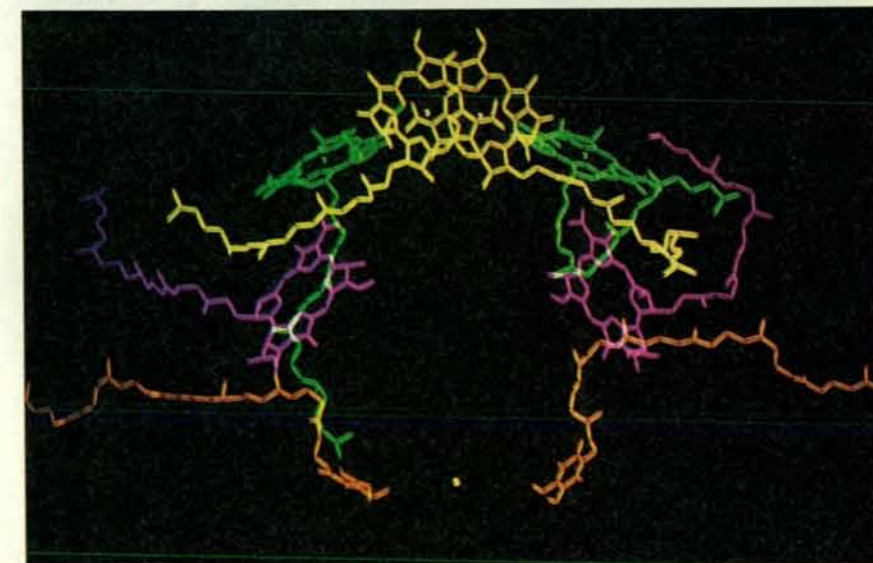
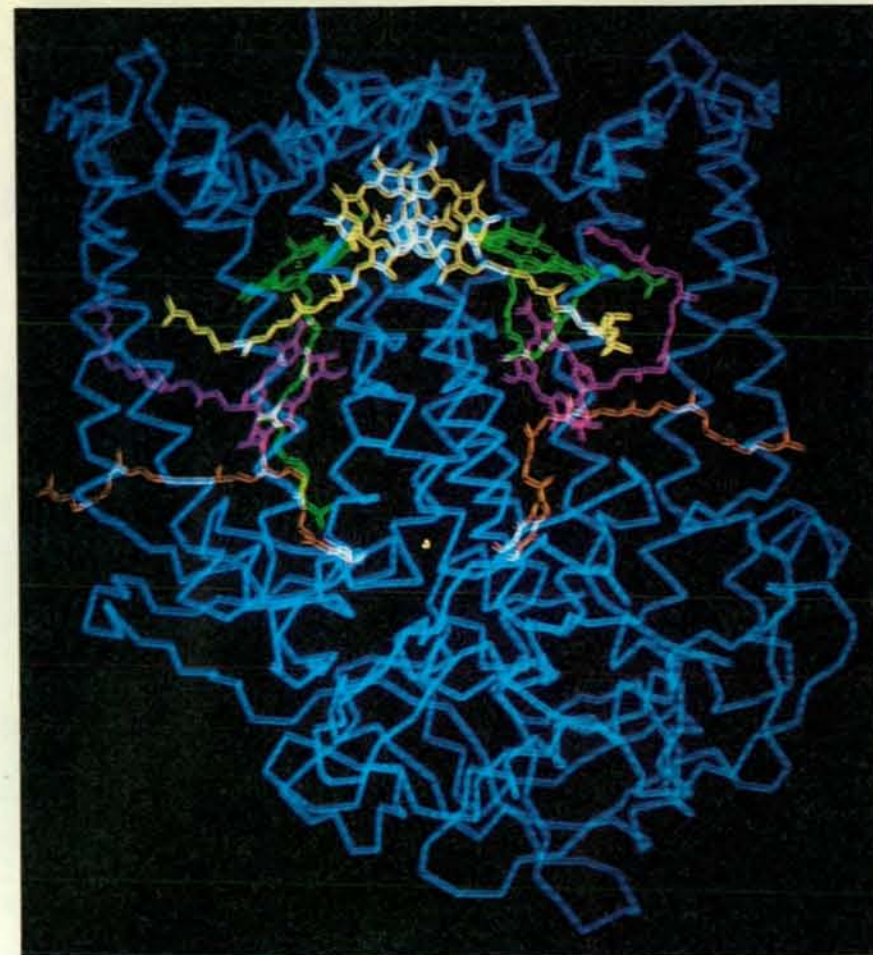
La spettroscopia di assorbimento può determinare, nell'arco di tempo di un picosecondo (cioè di 10^{-12} secondi), la collocazione temporale e la sequenza delle reazioni che costituiscono i primi stadi della fotosintesi. Due deboli fasci sonda (*al centro*) vengono fatti passare attraverso un campione contenuto in una celletta (*in primo piano*). Essi sono così deboli che non possono innescare una fotosintesi molto intensa. I fotorecettori (non visibili nella fotografia) determinano la composizione chimica del campione analizzando la luce che lo attraversa. Un intenso lampo di luce laser (*a destra*), di durata inferiore a un picosecondo, stimola la fotosintesi in una piccola regione attraversata anche da uno dei fasci sonda. Dato che il lampo è così breve, tutti i centri di reazione della fotosintesi (cioè i complessi molecolari entro cui si svolgono le prime reazioni della fotosintesi) svolgono un'attività fotosintetica simultanea. A mano a mano che la fotosintesi procede, la composizione chimica del campione si modifica e i cambiamenti possono essere determinati analizzando la luce del fascio sonda. L'apparecchiatura sperimentale che si vede nella fotografia si trova nel laboratorio di Michael R. Wasielewski dell'Argonne National Laboratory.

La spettroscopia ha rivelato la sequenza e la collocazione temporale delle reazioni di trasporto degli elettroni, che avvengono nel centro di reazione della fotosintesi dopo che questo ha assorbito un fotone. Che dire della struttura del centro di reazione? Quale via fisica segue l'elettrone passando da un lato all'altro del doppio strato lipidico?

A questi interrogativi può rispondere la cristallografia a raggi X, che attualmente è l'unico metodo in grado di rivelare la struttura su scala atomica di molecole biologiche complesse. La tecnica consiste nel colpire un cristallo con un fascio di raggi X e nel registrare le direzioni in cui esso diffrange i raggi. Dal diffrattogramma è possibile determinare la struttura molecolare del cristallo.

La principale difficoltà nell'applicare la cristallografia a raggi X all'analisi del centro di reazione della fotosintesi è che il campione da analizzare deve assumere la forma di un cristallo regolare. Ottenere un cristallo di questo tipo da una proteina è difficile e richiede molta cura e una buona dose di fortuna da parte del cristallografo. Negli ultimi vent'anni sono state cristallizzate e analizzate alcune centinaia di proteine idrosolubili, ma solo in epoca recentissima è stato trovato il modo di cristallizzare proteine di membrana non solubili in acqua, come quelle del centro di reazione della fotosintesi.

È stato possibile cristallizzare le proteine di membrana soprattutto per merito di recenti innovazioni tecniche nella progettazione e nell'impiego di alcune piccole molecole organiche con un'estremità idrofila (cioè che viene attratta dall'acqua) e un'estremità idrofoba (che viene respinta dall'acqua). Le estremità idrofobe di queste molecole possono legarsi alle regioni idrofobe delle proteine di membrana, in modo che le estremità idrofile delle molecole rimangano esposte. L'aggregato risultante di piccole molecole e di proteine può quindi essere disciolto in una soluzione acquosa e in seguito cristallizzato. Le molecole proteiche e le molecole più piccole cristallizzano insieme, in un processo noto come cocristallizzazione.

Nel 1983, Hartmut Michel e Johann Deisenhofer del Max Planck Institut für Biochemie a Martinsreid hanno applicato le tecniche sopra menzionate alla cristallizzazione delle proteine del centro di reazione di *Rhodopseudomonas* e alla determinazione della loro struttura. Il loro lavoro è stato un'impresa straordinaria: infatti, la proteina del centro di reazione di *Rhodopseudomonas* è tuttora l'unica proteina di membrana intera che sia nota con una risoluzione a livello atomico.

La struttura del centro di reazione della fotosintesi nei batteri, illustrata in queste immagini ottenute al calcolatore, è stata determinata mediante cristallografia ai raggi X. L'immagine in alto mostra l'intero centro di reazione della fotosintesi, in gran parte formato da un complesso proteico (*in blu*), quella in basso mostra solo le molecole prostetiche, che sono incluse nel complesso proteico. La proteina stessa è inserita in un doppio strato lipidico (qui non visibile), che delimita una struttura interna del batterio, la vescicola fotosintetica. La «coppia speciale» di molecole di clorofilla, che assorbe energia dai fotoni viene indicata in giallo. Altre molecole di clorofilla, dette «voyeur» perché si trovano vicine ai siti di reazione, ma non vi partecipano, sono indicate in verde, le molecole di feofitina in color magenta e quelle di chinone in arancione. La macchiolina gialla vicino alla base del centro di reazione rappresenta uno ione ferro. Le posizioni e gli orientamenti di alcune parti delle molecole prostetiche sono approssimati; le coordinate esatte non sono state ancora determinate. Le immagini sono state ottenute da Chong-Huan Chang, David M. Tiede, James R. Norris, Jr., e Marianne Schiffer dell'Argonne National Laboratory.

Nel complesso, le informazioni raccolte con gli studi cristallografici e le conoscenze acquisite dagli spettroscopisti danno un'immagine chiara, nello spazio e nel tempo, del comportamento chimico del centro di reazione della fotosintesi e delle reazioni che trasportano l'elettrone da una superficie all'altra del doppio strato lipidico.

Il principale elemento strutturale del centro di reazione è un grosso complesso proteico, incluso nel doppio strato lipidico. Alcune molecole più piccole, note come molecole prostetiche, o *helper* (cioè coadiutrici; fra di esse vi sono anche molecole di clorofilla), sono a loro volta parzialmente incluse nel complesso. Le molecole prostetiche costituiscono la via elettroconduttrice seguita dagli elettroni durante la fotosintesi.

Uno dei caratteri più sorprendenti di questa struttura sta nel fatto che possiede una simmetria di rotazione binaria quasi perfetta; in altre parole, un centro di reazione della fotosintesi, se venisse ruotato di 180 gradi attorno al proprio asse centrale (l'asse che decorre da una superficie del doppio strato lipidico all'altra), apparirebbe quasi identico a un centro di reazione non ruotato. Nel caso specifico, le molecole prostetiche sono disposte in due spirali, situate simmetricamente sui lati opposti della proteina centrale. Un'osservazione sperimentale curiosa e non ancora spiegata è che, malgrado questa simmetria, sembra che gli elettroni seguano solo una di queste spirali durante la fotosintesi.

Il processo fotosintetico inizia quando un fotone colpisce una coppia di molecole di clorofilla, chiamata «coppia speciale» e situata all'estremità del centro di reazione che è più vicina alla superficie interna del doppio strato lipidico. La coppia speciale si trova in corrispondenza della giunzione delle due spirali di molecole prostetiche. Un elettrone appartenente alla coppia speciale assorbe l'energia del fotone e si sposta su un gruppo prostetico vicino, una molecola di feofitina, che è molto simile alla clorofilla per composizione e struttura. Questo stadio del processo è estremamente rapido: impiega circa quattro bilionesimi di secondo. Durante questa reazione l'elettrone passa vicino, ma sembra senza congiungersi, a un'altra molecola di clorofilla (chiamata appropriatamente molecola «voyeur»). Sulla coppia speciale rimane un eccesso di carica positiva.

L'elettrone si muove, quindi, dalla feofitina a una molecola di chinone, che si trova all'estremità della catena a spirale di molecole prostetiche ed è vicina alla superficie esterna del doppio strato lipidico. Dalla molecola di chinone, l'elettrone passa attraverso la proteina centrale e raggiunge la molecola di chinone che si trova all'estremità dell'altra spirale di molecole prostetiche (la spirale lungo la quale non avviene la conduzione degli elettroni). Quest'ultimo stadio del processo è estremamente lento: si svolge a una velocità che è circa 10^8 volte inferiore a quella delle reazioni di trasporto dell'elettrone tra la coppia speciale e la molecola di feofitina.

Nel frattempo, una molecola globulare, idrosolubile, di citocromo cede un elettrone alla coppia speciale, assumendo quindi una carica positiva e neutralizzando la coppia speciale. Poi l'intero processo riprende daccapo: un altro fotone colpisce la coppia speciale e un altro elettrone procede lungo una spirale fino a una molecola di chinone e poi, attraverso la proteina centrale, giunge alla molecola di chinone che si trova all'estremità dell'altra spirale.

Questa seconda molecola di chinone, che ora porta due elettroni in più, si allontana dalla proteina del centro di reazione della fotosintesi per partecipare agli stadi successivi del processo fotosintetico, che si svolgono in corrispondenza della superficie esterna del doppio strato lipidico. Inoltre, una seconda molecola di citocromo cede un elettrone alla coppia speciale, neutralizzandola. La separazione delle cariche, che serve a immagazzinare l'energia dei fotoni, è ora completa: due molecole di citocromo, vicino alla superficie interna del doppio strato lipidico, hanno acquisito una carica positiva e due elettroni si sono spostati dalla superficie interna del doppio strato lipidico alla superficie esterna.

Il centro di reazione della fotosintesi dimostra di avere notevole efficienza nel catturare l'energia luminosa: cattura, infatti, l'energia del 98-100 per cento dei fotoni che assorbe. Se lo si considera una batteria, esso ha un rendimento del 50 per cento: l'energia immagazzinata nelle cariche separate è circa la metà di quella contenuta nei fotoni. Il resto dell'energia viene utilizzato nelle reazioni che trasportano gli elettroni lungo la catena di molecole prostetiche.

La spettroscopia e la cristallografia permettono di avere una visione approfondita del processo della fotosintesi in *Rhodopseudomonas*. Il terzo punto di vista da cui il processo è stato esaminato è quello della genetica molecolare.

Come abbiamo detto, è importantissimo per il genetista che *Rhodopseudomonas* possa sopravvivere senza svolgere attività fotosintetica. Ciò permette a mutanti con difetti genetici che riguardano i centri di reazione della fotosintesi di sopravvivere e di moltiplicarsi. Pertanto è possibile studiare in batteri viventi l'effetto di varie mutazioni.

Nell'analisi genetica di un organismo è importante non solo che i mutanti sopravvivano, ma anche che vi sia un modo di trasferire il DNA, la sostanza di cui sono fatti i geni, da un organismo all'altro per produrre incroci genetici. Geni alterati possono così essere inseriti in vari individui. Circa dieci anni fa, uno di noi (Marrs) scoprì che una specie di *Rhodopseudomonas*, *R. capsulata*, possiede un meccanismo del genere. Capita a volte che il materiale genetico di questo batterio si frammenti in sequenze lunghe circa cinque geni ciascuna. I siti dove avviene la frammentazione sembrano essere casuali e variano da un individuo all'altro. Ogni frammento di materiale genetico viene poi incapsulato in una piccola particella che sembra un virus. La parete cellulare del batterio quindi esplode, rilasciando questi agenti per il trasporto di geni, che finiscono su altri batteri *R. capsulata*, inserendovi il loro breve filamento di materiale genetico. I geni così trasportati si integrano, quindi, nel DNA dei batteri riceventi.

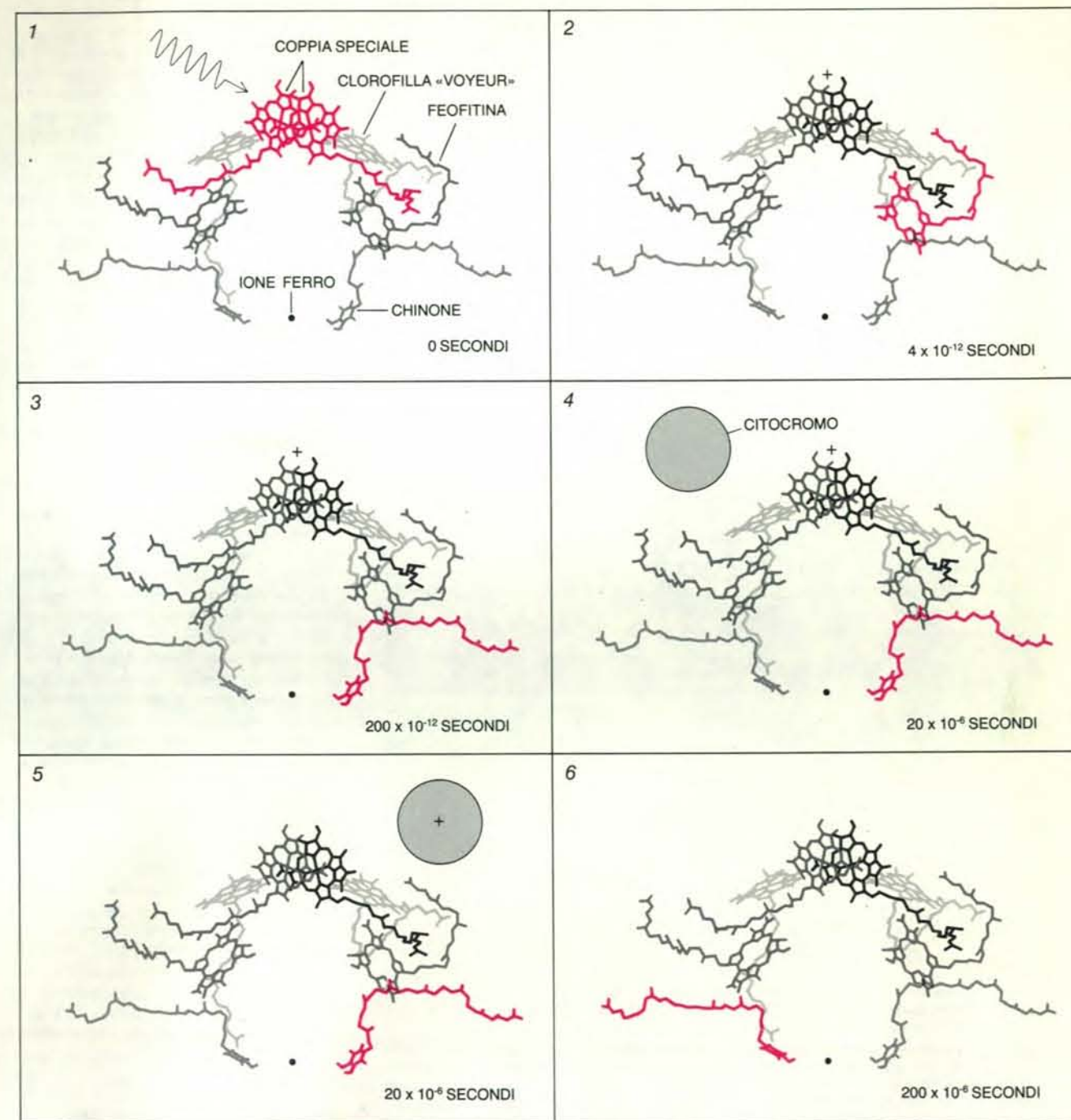
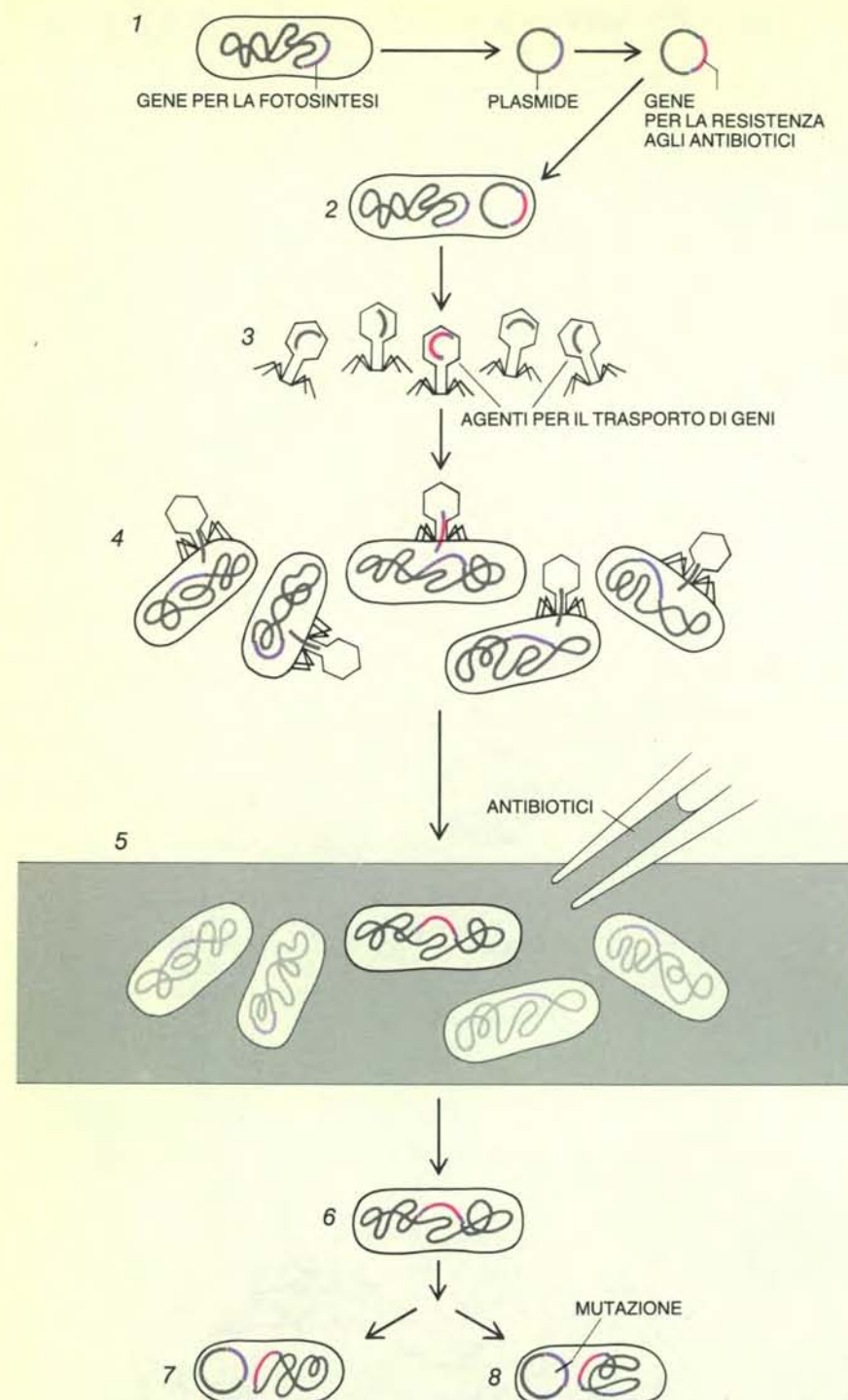
Gli agenti per il trasporto di geni sono utili per produrre incroci genetici e possono anche servire a costruire la mappa dei geni dei batteri. Notando con quale frequenza due geni diversi sono incapsulati nello stesso agente di trasporto, il ricercatore può riuscire a stimare quanto vicini essi si trovino sul cromosoma batterico (cioè il filamento di DNA che contiene la maggior parte dei geni batterici). Quanto più spesso due geni sono incapsulati insieme, tanto più vicini essi devono trovarsi su quel cromosoma.

I batteri della specie *R. capsulata* possono anche scambiarsi tra loro materiale genetico mediante «accoppiamento». Questo processo avviene per mezzo di un'appendice, il pilo, che permette di trasferire il DNA da un batterio «maschio» a un batterio «femmina». Il comportamento offre un altro modo di produrre incroci genetici.

Per mezzo di questi incroci, uno di noi (Marrs) è riuscito a isolare, nel 1980, un frammento di DNA che contiene la maggior parte dei geni necessari alle reazioni alla luce della fotosintesi. Questo frammento include i geni per tutte le proteine che si trovano nel centro di reazione della fotosintesi e per le proteine che costituiscono una struttura separata, nota come complesso «antenna», per la raccolta della luce. Infatti questa antenna serve a catturare l'energia dei fotoni e a incanalarla nel centro di reazione. E' stata questa la prima volta in cui, in un qualsiasi organismo fotosintetizzante, è stato isolato il DNA di un centro di reazione della fotosintesi.

Sulla base di questa ricerca, uno di noi (Youvan) ha caratterizzato il DNA isolato. Con i suoi collaboratori del Lawrence Berkeley Laboratory, ha localizzato ogni gene all'interno del frammento e ha determinato per ciascuno la sequenza nucleotidica (cioè la successione di quegli elementi che determinano la costituzione specifica della proteina codificata dal gene). La sequenza nucleotidica è stata la base di un'analisi al calcolatore della struttura della proteina: un'analisi che ha previsto aspetti della struttura globale del centro di reazione diversi anni prima che essa venisse ricostruita mediante cristallografia a raggi X. La conoscenza della sequenza nucleotidica del centro di reazione della fotosintesi di un batterio ha anche permesso di confrontarla con la sequenza corrispondente, presente nelle piante superiori. La stretta similitudine tra le due sequenze sta a indicare che la costituzione delle strutture fotosintetiche fondamentali è quasi universale.

La conoscenza della costituzione genetica del centro di reazione della fotosintesi nel batterio *R. capsulata* fornisce gli strumenti necessari per affrontare una questione della massima importanza: in quali modi specifici i vari elementi del centro di reazione contribuiscono al funzionamento globale di quest'ultimo?

Un modo per dare una risposta a tale domanda consiste nel produrre mutazioni semplici nei geni che codificano per le strutture del centro di reazione. Gli organismi che portano le mutazioni possono quindi essere osservati con tecniche

Le reazioni chimiche responsabili dei primi stadi della fotosintesi trasportano un elettrone a un'estremità del centro di reazione, lasciando una regione con carica elettrica positiva all'altra estremità. Nel primo stadio della fotosintesi (*1*), un fotone viene assorbito dalla «coppia speciale» di molecole di clorofilla e trasferisce la sua energia a un elettrone presente in essa. (In ogni stadio, le molecole con un elettrone eccitato sono in rosso.) L'elettrone si sposta quindi (*2*) su una molecola di feofitina, oltrepassando una molecola di clorofilla «voyeur» e lasciando una carica positiva sulla coppia speciale. L'elettrone procede quindi (*3*) verso la molecola di chinone, all'estremità di una catena di molecole prostetiche in forma di spirale. In questo stadio, una molecola di citocromo (di forma globulare), che si sposta liberamente nella soluzione, si avvicina alla coppia speciale (*4*) e le cede un elettrone. Il citocromo acquista così una carica positiva (*5*) e le molecole della coppia speciale tornano a essere neutre. Quindi, l'elettrone eccitato, che si era portato su una molecola di chinone, passa a una seconda molecola di chinone (*6*). La separazione dell'elettrone e l'instaurarsi di una regione con carica positiva costituiscono un immagazzinamento di energia.

Le tecniche della genetica molecolare permettono di sostituire geni convenzionali con geni mutati e di produrre così un batterio che manca del gene per il centro di reazione della fotosintesi; è anche possibile verificare nel batterio la funzione dei geni alterati. Il gene per la fotosintesi (*in colore viola*), prelevato da un batterio normale (*1*), viene inserito in un plasmide (un piccolo anello di DNA). Esso viene poi eliminato, ma i segmenti di DNA che lo fiancheggiavano sono lasciati *in situ*. Al suo posto viene inserito un gene per la resistenza agli antibiotici (*in rosso*). Il plasmide è poi introdotto in un batterio normale (*2*). Questo frammenta il proprio DNA in piccoli segmenti, che incapsula in particelle simili a virus, chiamate agenti per il trasporto di geni (*3*). Questi agenti si attaccano ad altri batteri (*4*), inoculando in essi il DNA che trasportano. Il gene per la resistenza agli antibiotici, essendo intercalato tra segmenti di DNA che normalmente fiancheggiano il gene per la fotosintesi, va a sostituire proprio questo gene nel DNA batterico. I batteri sono quindi messi in presenza di antibiotici (*5*), che impediscono agli individui senza il nuovo gene di replicarsi. Gli organismi superstiti non hanno il gene per la fotosintesi (*6*). Quando un plasmide contenente il gene normale per la fotosintesi viene inserito in essi (*7*), la loro attività fotosintetica procede normalmente; batteri di questo genere confermano che le manipolazioni genetiche non hanno danneggiato in modo irreversibile la capacità fotosintetica. E' anche possibile inserire un plasmide contenente un gene per la fotosintesi che comprenda una mutazione (*8*). Verificando la funzione fotosintetica dei batteri risultanti si può rivelare l'effetto del gene alterato.

spettroscopiche, per determinare in che modo la loro attività fotosintetica differisca da quella dei batteri non mutati. Esperimenti di questo tipo stanno già fornendo informazioni basilari su certi aspetti della fotosintesi batterica. In alcuni casi l'informazione acquisita si può applicare anche alle piante superiori.

Per esempio, la molecola di chinone che al termine del processo riceve gli elettroni in *Rhodopseudomonas* funziona circa allo stesso modo, e in un ambiente simile, nelle piante superiori. La struttura dell'incavo proteico nel centro di reazione, a cui si lega, nei batteri, la molecola di chinone, è stata determinata di recente mediante i raggi X da Feher e collaboratori. Uno di noi (Youvan), insieme a un laureando (Edward J. Bylina), ha introdotto alcune mutazioni in questo incavo. Parte dei batteri che ne sono risultati non avevano affatto attività fotosintetica, parte avevano una funzione fotosintetica alterata e parte, infine, avevano una funzione fotosintetica normale. La misura in cui la fotosintesi è alterata dipende dalla specifica mutazione. In alcuni ceppi mutati, la funzione della molecola di chinone è modificata in modo tale che questi ceppi risultano resistenti a certi erbicidi (come l'atrazina) che uccidono le piante inibendo il funzionamento del chinone.

Quest'ultima informazione, che deriva dallo studio dei batteri, potrà essere presto applicata alle piante superiori. Nel caso della resistenza agli erbicidi, analoghe mutazioni potrebbero essere prodotte nei fotosistemi di piante superiori, come la soia, che non sono resistenti all'atrazina. Piante di soia rese resistenti all'atrazina attraverso mutazioni potrebbero essere messe a dimora in campi nei quali erano cresciute in precedenza colture, come il granturco, che sono naturalmente resistenti all'atrazina e spesso vengono irrorate abbondantemente con questo erbicida al fine di eliminare le piante infestanti.

A livello di ricerca di base, gli studi su *Rhodopseudomonas* possono condurre a scoperte che riguardano le proprietà elettroniche di base di varie proteine. Un ricercatore, mediante manipolazioni genetiche che introdurrebbero centri di reazione alterati al posto di quelli naturali, potrebbe riuscire a scoprire con quanta rapidità o con quanta efficienza le nuove proteine conducano gli elettroni. Si possono prevedere molti test sperimentali simili.

La conoscenza approfondita e particolareggiata che è alla base di questi esperimenti è stata raggiunta grazie all'unione di tre campi di ricerca che sono stati spesso considerati quasi indipendenti l'uno dall'altro. Ciascun gruppo di ricercatori, esaminando il problema da una diversa angolazione, ha chiarito un differente aspetto della fotosintesi, così che ora è possibile avere una visione completa dell'intero processo.

Rivelatori di onde gravitazionali

Operativi entro gli inizi del prossimo decennio, potrebbero rivelare emissioni provenienti da sorgenti extragalattiche e fornire in questo modo una conferma alle previsioni della teoria della relatività generale

di Andrew D. Jeffries, Peter R. Saulson, Robert E. Spero e Michael E. Zucker

Secondo la teoria della relatività generale proposta da Albert Einstein, la Terra sarebbe continuamente investita da radiazioni energetiche prodotte da interazioni gravitazionali di stelle lontane. L'energia emessa da una perturbazione cosmica, come l'esplosione di una stella, viaggerebbe alla velocità della luce sotto forma di onde gravitazionali che, nel loro moto, provocherebbero distorsioni nelle regioni di spazio attraversate. Si potrebbero rilevare tali onde misurando i moti di masse di controllo accuratamente isolate.

Purtroppo, queste perturbazioni sono troppo deboli per essere evidenziate da strumenti comuni e i rivelatori di onde gravitazionali messi a punto negli ultimi 25 anni non sono ancora riusciti a registrare il passaggio di un'onda. Potrebbe apparire sorprendente la difficoltà di rilevare onde gravitazionali, considerando quanto sia familiare la gravità statica in esperienze quotidiane come la caduta di un oggetto o il moto orbitale dei pianeti; tuttavia, le masse coinvolte in questi fenomeni non sono abbastanza grandi e veloci da generare onde apprezzabili. È probabile che i primi segnali che potremo rilevare giungano da corpi celesti lontani più massicci del Sole e in moto a una velocità di poco inferiore a quella della luce.

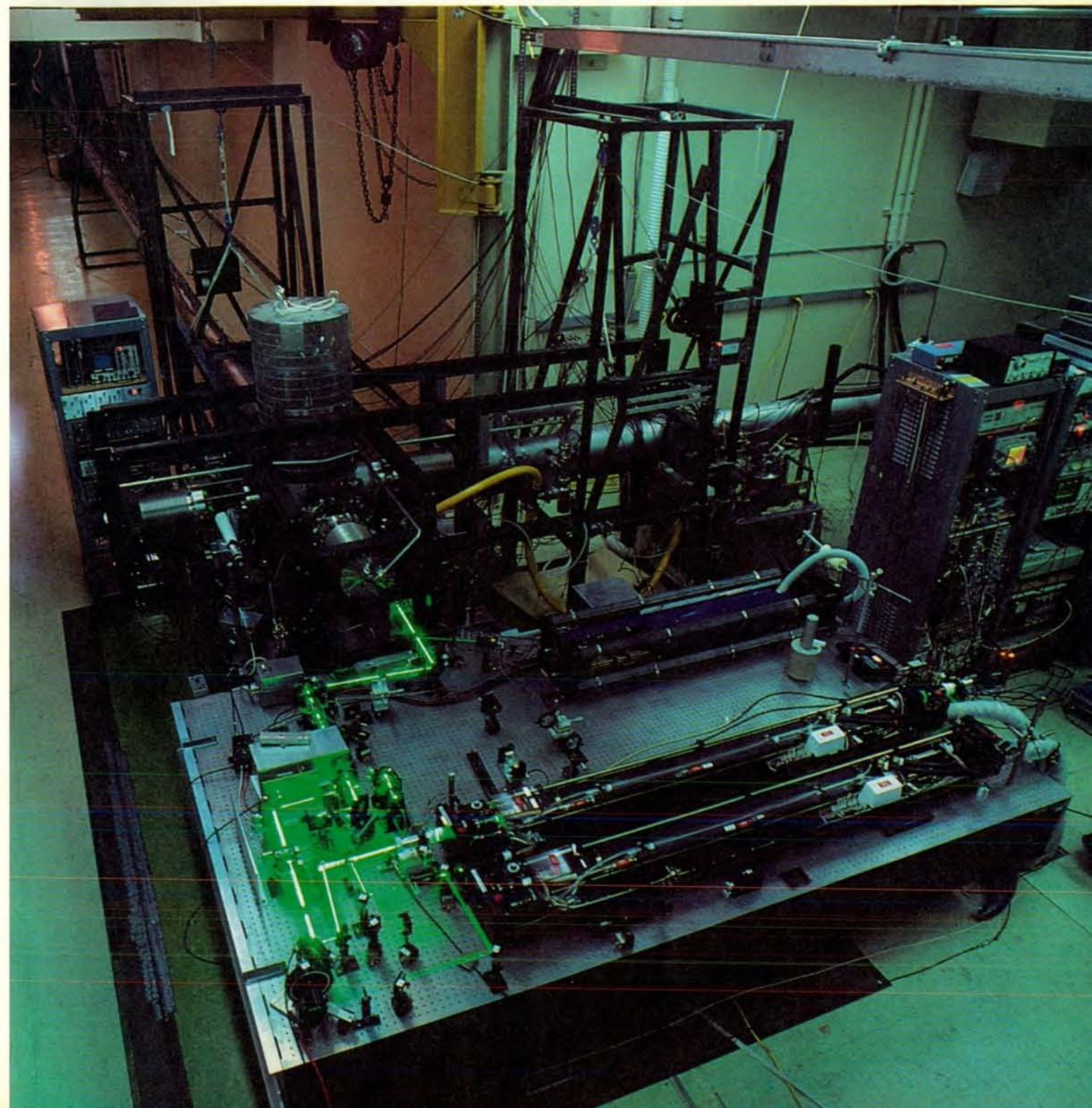
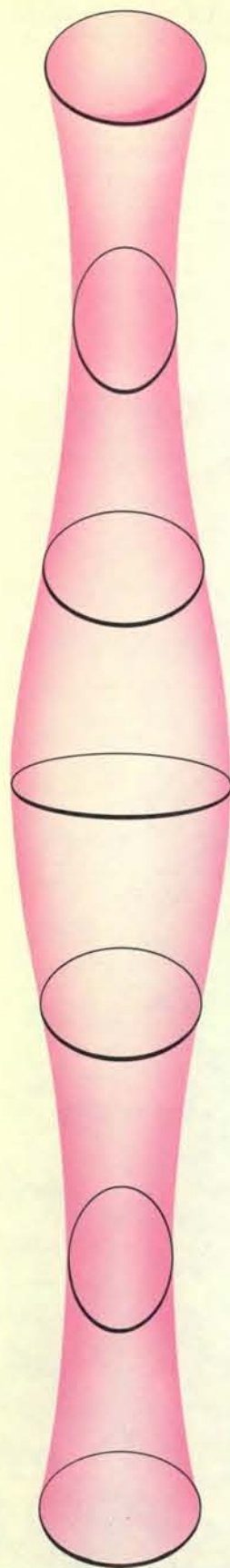
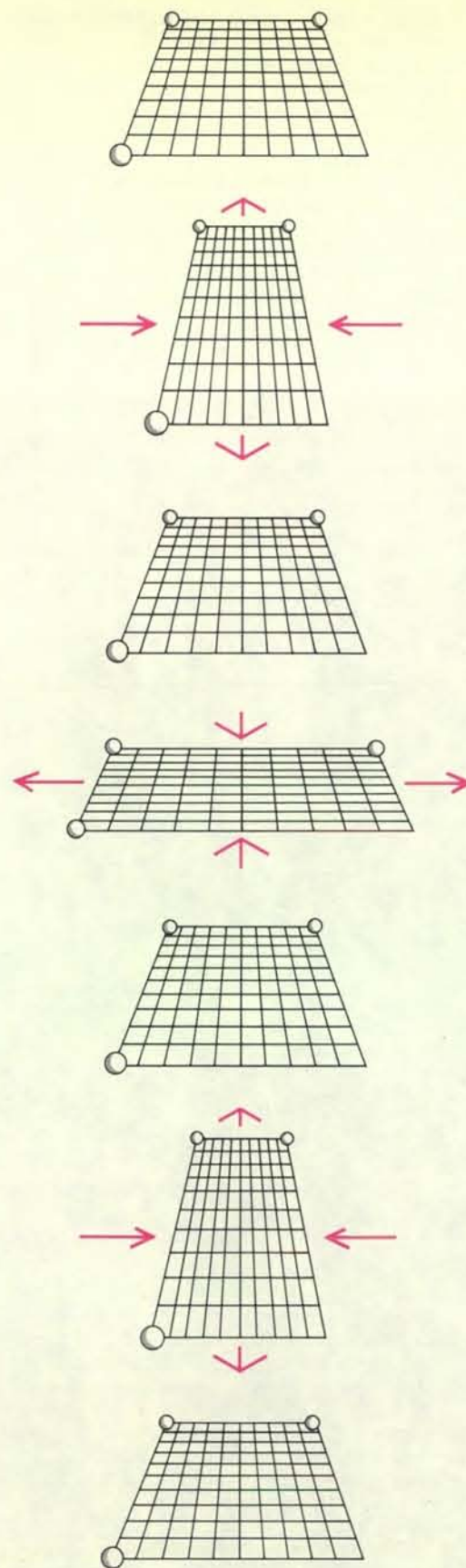
Attualmente, si sta mettendo a punto una nuova serie di esperimenti più sensibili per rilevare le onde gravitazionali e gli osservatori ancora in fase di progetto potrebbero divenire operativi all'inizio degli anni novanta. Prototipi di rivelatori basati sull'interferometria laser sono già stati costruiti e sperimentati in laboratori europei e statunitensi. Negli Stati Uniti, gli autori stanno lavorando nei gruppi diretti da Ronald W. P. Drever del California Institute of Technology e da Rainer Weiss del Massachusetts Institute of Technology alla costruzione di una coppia di grandi interferometri situati alle estremità opposte del continente, uno nel deserto di Mojave, nella California meridionale, e l'altro a Columbia, nel Maine. I due strumenti saranno collegati elettronicamente per potere operare come un unico osservatorio, chiamato osservatorio a interferometria laser di onde gravitazionali, o LIGO. Ogni strumento sarà 1000 volte più sensibile del miglior rivelatore oggi esistente, ma il costo totale della coppia (circa 60 milioni di dollari) non supererà quello di un grande telescopio ottico.

Possiamo aspettarci che LIGO e altri osservatori del genere ci rivelino un'immagine dell'universo nuova e sotto molti aspetti diversa da quella che vediamo grazie all'astronomia convenzionale. Fino agli anni trenta, l'unico mezzo che permettesse di studiare l'universo lontano era la radiazione elettromagnetica alle frequenze della luce visibile. L'immagine dell'universo fu ampliata radicalmente con l'avvento della radioastronomia e, successivamente, con l'aprirsi delle «finestre» infrarossa, X e gamma. Benché ogni forma di radiazione elettromagnetica ci riveli fenomeni particolari, le onde gravitazionali trasportano energia di tipo del tutto differente e quindi informazioni non ottenibili in altro modo sulle loro sorgenti. Gli osservatori di onde gravitazionali potrebbero trasformare completamente la nostra visione dell'universo.

La dimostrazione dell'esistenza di onde gravitazionali costituirebbe inoltre una verifica diretta delle previsioni della relatività generale. Un postulato fondamentale della relatività dice che nessun segnale può viaggiare a una velocità superiore a quella della luce. Nella teoria classica di «azione a distanza» della gravità, invece, una variazione di un campo gravitazionale si propagherebbe a velocità infinita. La semplice verifica che le onde gravitazionali si propagano alla velocità della luce costituirebbe un forte argomento a sostegno della teoria di Einstein.

Per capire l'influenza delle onde gravitazionali sulla materia, supponiamo che un'onda di questo tipo viaggi nello spazio. Immaginiamo di porre un tubo di gomma lungo e flessibile, di sezione circolare, nella direzione in cui l'onda si sposta. Muovendosi lungo il tubo, l'onda ne distorce la sezione in un'ellisse; l'asse maggiore aumenta di una frazione del suo diametro originario, mentre il minore diminuisce della stessa quantità. Dopo un semiperiodo, gli assi si invertono e sulla superficie del tubo si formano ondulazioni. Quando l'onda è passata, il tubo ritorna alla sua forma originale.

Quale effetto potrebbe avere un'onda di questo genere su una successione di masse ugualmente distanziate sospese in un piano? Supponiamo che il piano sia orizzontale e che l'onda si muova perpendicolarmente a esso. Mentre attraversa il piano, essa dapprima farà aumentare la distanza fra le masse in una direzione, per esempio nord-sud, e diminuire la distanza nella direzione perpendicolare, in questo caso est-ovest. Questa minuscola variazione nella separazione delle masse si può definire come l'ampiezza della deformazione prodotta dall'onda. Mentre questa continua a propagarsi attraverso il piano, gli stiramenti e le compressioni si invertiranno periodicamente: la separazione fra le masse nelle direzioni nord-sud ed est-ovest aumenterà e diminuirà alternativamente. Gli effetti prodotti dalle onde sono perpendicolari alla loro direzione di propagazione, perciò un'onda che si muove verso ovest influirà sulla separazione delle masse lungo l'asse nord-sud, ma non lungo quello est-ovest.

La produzione di onde gravitazionali è simile per molti aspetti alla produzione di onde elettromagnetiche. Le onde gravitazionali differiscono dalla più nota attrazione di gravità nello stesso modo in cui le onde luminose e radio differiscono dall'elettricità statica e dal magnetismo. Un corpo in movimento dotato di carica elettrica irradia onde elettromagnetiche di ampiezza proporzionale alla sua cari-

Il rivelatore di onde gravitazionali del California Institute of Technology, prototipo dei grandi osservatori la cui entrata in funzione è prevista per il prossimo decennio, comprende un interferometro laser contenuto in due tubi sotto vuoto perpendicolari fra loro, lunghi 40 metri. Il fascio laser in primo piano è diretto da specchi e fibre ottiche in un contenitore sotto vuoto (*in secondo piano a sinistra*), in cui si trova un divisore di fascio (uno specchio parzialmente riflettente) che separa il fascio in parti identiche fra i due tubi. Specchi montati su coppie di masse appese al termine dei tubi riflettono la luce. I fasci luminosi vengono riflessi avanti e indietro per circa 10 000 volte, e si osserva poi l'interferenza risultante. Il passaggio di un'onda gravitazionale farebbe variare un poco la distanza fra le masse di una o di entrambe le coppie e quindi l'interferenza. L'apparecchio è sensibile a variazioni fino a 3×10^{-16} metri (un terzo del diametro di un protone), che durino solo un millisecondo. Il Massachusetts Institute of Technology possiede un interferometro da 1,5 metri, mentre un altro da cinque metri è in costruzione.

Le onde gravitazionali dovrebbero provocare distorsioni nelle regioni dello spazio attraversate. Gli effetti vengono mostrati nel caso di un tubo di gomma liberamente fluttuante (*a sinistra*) e di masse disposte in un piano (*a destra*). In questo caso, le onde si muovono verticalmente.

ca e alla sua accelerazione. La carica gravitazionale di un corpo è la sua massa; perciò l'ampiezza di un'onda gravitazionale dovrebbe essere proporzionale alla massa e all'accelerazione del corpo. Ma il principio di conservazione della quantità di moto richiede che un sistema isolato non abbia accelerazione netta; per ogni azione vi è una reazione uguale e contraria. Perciò, la maggior parte delle onde emesse da una massa che accelera è annullata dalle onde emesse da una massa che simultaneamente retrocede. L'annullamento tuttavia non è perfetto in quanto le due masse non hanno posizioni esattamente coincidenti. Questo tenue spostamento nella collocazione delle masse fa sì che le onde siano leggermente sfasate e quindi che parte della radiazione gravitazionale possa sfuggire dal sistema.

La quantità di radiazione gravitazionale emessa da un corpo dipende dalla disomogeneità nella distribuzione della sua massa. Il momento di quadrupolo è una misura della disomogeneità; un pallone da calcio ha momento di quadrupolo nullo, mentre un pallone da rugby ne ha uno molto grande. Onde gravitazionali intense vengono emesse quando il momento di quadrupolo di un oggetto di grande massa subisce rapide variazioni. La massa e la velocità dell'oggetto, tuttavia, devono essere enormi; una barra d'acciaio di 500 tonnellate, in rotazione così rapida da sfiorare il limite di disintegrazione, genererebbe onde gravitazionali in grado di provocare, in un sistema di controllo, deformazioni pari, al massimo, a una parte su 10^{40}. Una simile variazione è ovviamente troppo piccola per essere rilevata.

Le stelle e altri corpi celesti dovrebbero generare onde gravitazionali molto più intense di quelle emesse da qualsiasi sorgente terrestre. Un sistema stellare binario, formato da due stelle che orbitano intorno a un comune centro di massa, dovrebbe produrre onde gravitazionali continue con un periodo fondamentale pari a metà del periodo orbitale (dato che il momento di quadrupolo assume lo stesso valore due volte per ogni orbita). Quasi metà delle stelle conosciute si trovano in sistemi binari. Questi sistemi hanno un momento di quadrupolo elevato e variabile, ed è possibile prevedere con precisione la quantità di radiazione gravitazionale da essi generata. Le onde gravitazionali più intense emesse da un sistema binario dovrebbero produrre una deformazione in un rivelatore sulla Terra pari a una parte su 10^{20}; maggiore quindi di 20 ordini di grandezza rispetto alla deformazione prodotta dalla barra d'acciaio in rotazione.

Una piccola parte dei sistemi binari conosciuti è composta da due stelle di neutroni in orbita stretta. Una stella di neutroni è un corpo celeste di massa simile a quella del Sole e con un raggio di una decina di chilometri soltanto, rilevabile talvolta sotto forma di una pulsar: una radiosorgente che, ruotando su se stessa, proietta tutt'intorno un fascio di onde radio, che appare a intermittenza, come quello luminoso di un faro mentre spazza la regione circostante.

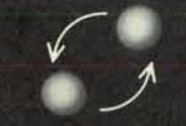
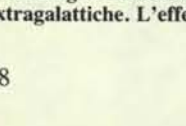
Nel 1974 Joseph H. Taylor, Jr., e Russell A. Hulse, allora all'Università del Massachusetts ad Amherst, scoprirono, in un sistema binario, una pulsar con un periodo orbitale di otto ore. Benché solo una delle stelle sia una pulsar, lo studio dei parametri orbitali indica che anche la compagna è una stella di neutroni. Misurando esattamente i tempi di arrivo dei segnali provenienti dalla pulsar, Taylor e collaboratori hanno potuto calcolare con grande precisione le orbite delle due stelle di neutroni. Si è così visto che il periodo diminuisce gradualmente, il che indica che le due stelle stanno avvicinandosi l'una all'altra su traiettorie a spirale. La perdita di energia osservata è in ottimo accordo con quella calcolata per effetto della radiazione gravitazionale; ossia, sembra che le onde gravitazionali estraggano dal sistema una quantità di energia e di momento angolare sufficiente a giustificare esattamente l'effetto che si osserva. La previsione teorica si accorda ai risultati delle misurazioni con uno scarto inferiore all'uno per cento. Questa notevole concordanza tra teoria ed esperimento fornisce una solida prova dell'esistenza della radiazione gravitazionale.

La prima dimostrazione dell'esistenza di onde gravitazionali proviene da una pulsar binaria scoperta nel 1974 da Joseph H. Taylor, Jr., e Russell A. Hulse, allora all'Università del Massachusetts ad Amherst. Una pulsar binaria consiste di due stelle che orbitano intorno al comune centro di massa (*in alto*): una di esse è una pulsar, una radiosorgente che sembra emettere impulsi intermittenti. Il periodo orbitale del sistema, che attualmente è di otto ore, sta diminuendo lentamente (*in basso*). Tale diminuzione, calcolata misurando con la massima esattezza i tempi di arrivo dei segnali radio, implica che le due stelle stiano perdendo gradualmente energia e avvicinandosi l'una all'altra su traiettorie a spirale. La perdita di energia osservata (*punti in nero*) si accorda molto bene con quella prevista in base a calcoli di emissione di radiazione gravitazionale (*curva in colore*); sembra che le onde gravitazionali estraggano dal sistema la quantità esatta di energia e di momento angolare che permette di spiegare le osservazioni.

Quando un sistema binario muore, le stelle che lo compongono cadono rapidamente verso il centro seguendo una traiettoria a spirale, fino a che collidono o si disintegrano. Nel caso delle stelle di neutroni, entrambi gli eventi produrrebbero un intenso impulso di onde gravitazionali. Il sistema sopra citato non farà questa fine per altri 100 milioni di anni circa, ma galassie come la nostra contengono un numero sufficiente di sistemi binari perché in ogni millennio si abbiano numerose collisioni tra coppie di stelle. Tuttavia, un millennio è un tempo di attesa molto lungo. LIGO è stato progettato con una sensibilità tale da rilevare segnali provenienti dall'ammasso della Vergine e da altri ammassi di galassie vicini, e potrà perciò individuare diversi eventi all'anno in grado di produrre una deformazione di una parte su 10^{22} nel rivelatore.

Anche la nascita di una stella di neutroni dovrebbe essere annunciata da un intenso impulso di onde gravitazionali. Quando una stella massiccia muore in una grande esplosione di supernova, il suo nucleo collassa rapidamente e forma una stella di neutroni. Sebbene i particolari del collasso non siano ancora noti con precisione, è possibile supporre che una quantità pari anche allo 0,1 per cento della massa della stella di neutroni venga trasformata in onde gravitazionali. Un'esplosione di supernova nella Galassia, un evento che accade circa una volta ogni 30 anni, potrebbe provocare una deformazione di una parte su 10^{18}. La recente supernova nella Grande Nube di Magellano, una galassia satellite della Via Lattea, dovrebbe aver prodotto onde gravitazionali in grado di causare una deformazione di una parte su 10^{19}, appena al di sotto della soglia di sensibilità dei migliori rivelatori attuali. Le numerose supernove che appaiono ogni anno nell'ammasso della Vergine dovrebbero causare deformazioni dell'ordine di una parte su 10^{21}.

Il rilevamento di onde gravitazionali provenienti da una supernova avrebbe profonde conseguenze. Sarebbe interessante misurare quanto tempo passa tra l'individuazione delle onde gravitazionali generate nel collasso del nucleo di una supernova e l'arrivo della radiazione luminosa emessa dagli strati esterni della stella che esplode. Se le onde gravitazionali e quelle luminose venissero rilevate simultaneamente, avremmo una conferma diretta della previsione relativistica che le onde gravitazionali si propagano alla velocità della luce. Se, per esempio, una supernova nell'ammasso della Vergine venisse osservata con i telescopi ottici entro un giorno dal rilevamento dell'impulso gravitazionale, le velocità dei due segnali corrisponderebbero con uno scarto di una parte su 10 miliardi.

Le onde gravitazionali generate da una supernova potrebbero fornire inoltre una nuova finestra di osservazione. Le radiazioni elettromagnetiche prodotte nell'interno di un nucleo che collassa sono bloccate dagli strati esterni della stella, che nascondono alla vista le fasi

più violente dell'esplosione. Il rilevamento di un impulso di neutrini dalla supernova nella Grande Nube di Magellano ha fornito un valido argomento a favore del modello del collasso del nucleo, ma persino i neutrini subiscono migliaia di eventi di diffusione mentre sfuggono dalla stella. Le onde gravitazionali, che interagisono così debolmente con la materia da poter attraversare senza attenuazioni l'atmosfera di una stella, potrebbero svelare i dettagli più fini del collasso. La radiazione gravitazionale potrebbe permetterci anche di vedere oggetti oscurati da polvere interstellare, come il centro della Galassia.

Una stella di neutroni giunta alla maturità può essere anch'essa una sorgente di onde gravitazionali se il suo asse di rotazione non è perfettamente simmetrico. Se, per esempio, lo strato più esterno della stella si fosse consolidato in una forma appiattita a causa della veloce rotazione, l'incontro con un'altra stella potrebbe spostare la posizione dell'asse. L'oscillazione risultante dovrebbe essere sufficientemente ampia da produrre onde gravitazionali molto intense.

Lo stesso effetto dovrebbe prodursi al momento della formazione di un buco nero o nella collisione di due buchi neri. I buchi neri sono oggetti estremamente densi, la cui esistenza è prevista dalla teoria della relatività generale, ma non è stata ancora definitivamente provata. Si pensa che un buco nero si formi quando una stella molto massiccia, dopo aver esaurito il combustibile nucleare, collassa sotto il proprio peso in un residuo estremamente denso che conserva l'intensa forza di gravità della stella, tanto intensa che neppure la luce può sfuggire dal «buco». Se buona parte della massa dell'universo fosse contenuta nei buchi neri, le collisioni fra di essi dovrebbero essere abbastanza frequenti perché i rivelatori ne possano individuare un certo numero. Si ritiene che alcuni sistemi stellari binari da cui provengono raggi X contengano un buco nero, anche se è impossibile provarlo formalmente senza un'opportuna sonda gravitazionale. Gli osservatori di onde gravitazionali potrebbero fungere adeguatamente da sonda, perché la radiazione gravitazionale emessa da un buco nero dovrebbe avere una «firma» caratteristica. Il rilevamento di tali segnali fornirebbe l'unica dimostrazione non ambigua dell'esistenza di oggetti con queste peculiari caratteristiche gravitazionali.

Un'altra possibile fonte di onde gravitazionali è il big bang: l'esplosione primordiale che tra 15 e 20 miliardi di anni fa diede origine all'universo. Alcune delle informazioni più importanti sull'universo primordiale ci vengono dall'osservazione del fondo cosmico di microonde, il resto della radiazione termica che pervadeva l'universo ai suoi inizi. Il rilevamento di un fondo cosmico di onde gravitazionali svelerebbe nuovi aspetti del big bang. Per esempio, eventuali fluttuazioni quantistiche avvenute nei primi 10^{-43} secondi di vita dell'universo potrebbero aver lasciato tracce tuttora osservabili. Anche le violente fasi iniziali della formazione delle galassie potrebbero essere svelate da onde gravitazionali residue.

SORGENTE	TIPO DI SEGNALE	FREQUENZA	INTENSITÀ
SISTEMA BINARIO	PERIODICO	1 MEGAHERTZ O MENO	10^{-21}
SISTEMA BINARIO CON STELLE DI NEUTRONI	QUASIPERIODICO	CONTINUO FINO A 1 CHILOHERTZ	10^{-22}
STELLA DI NEUTRONI IN ACCRESCIMENTO	PERIODICO	DA 200 A 800 HERTZ	3×10^{-27}
SUPERNOVA DI TIPO II	PULSATO	1 CHILOHERTZ	10^{-21}
BUCO NERO IN VIBRAZIONE	SINUSOIDALE APPIATTITO	10 CHILOHERTZ PER UNA MASSA SOLARE 10 HERTZ PER 1000 MASSE SOLARI	?
FORMAZIONE DI GALASSIE	RUMORE	BANDA LARGA 1 CICLO ALL'ANNO 300 HERTZ	10^{-14} 10^{-24}
BIG BANG	RUMORE	?	?

Le onde gravitazionali possono venire da molte sorgenti, galattiche ed extragalattiche. L'effetto della loro interazione con la Terra è minimo: per esempio, onde provenienti da un sistema binario stellare sposterebbero le masse di un rivelatore di 10^{-21} metri per metro di separazione.

Le onde gravitazionali prodotte nelle situazioni descritte avrebbero un impatto estremamente debole ed effimero quando investono la Terra. Le masse dei rivelatori verrebbero appena sfiorate, con uno spostamento nelle loro posizioni di appena 10^{-21} metri (un milionesimo del diametro di un protone) per metro di separazione; onde pulsate, provenienti per esempio da una supernova, durerebbero non più di qualche millisecondo. Come è possibile rilevare segnali così deboli?

I primi rivelatori di onde gravitazionali furono realizzati all'inizio degli anni sessanta da Joseph Weber dell'Università del Maryland a College Park. I suoi strumenti erano costituiti da una sola, massiccia barra di metallo in forma di un cilindro pieno, solitamente di alluminio. Rivelatori di questo tipo vengono costruiti in modo tale da vibrare in risposta alle onde gravitazionali. I migliori rivelatori a barra attuali sono sensibili a deformazioni fino a una parte su 10^{18}; tuttavia nessun segnale di questo livello è mai stato rilevato.

I tentativi di fare misurazioni più accurate con i rivelatori a barra si scontrano con alcune difficoltà; una di esse è che la sensibilità di un rivelatore di onde gravitazionali aumenta all'aumentare della lunghezza, ma non è possibile, in pratica, costruire barre più lunghe di qualche decina di metri. I nostri progetti prevedono la costruzione di rivelatori a interferometria della lunghezza di alcuni chilometri, costituiti da masse separate interconnesse solo da fasci di luce.

Il fenomeno dell'interferenza è associato con tutti i tipi di onde, comprese quelle elettromagnetiche. Le onde nell'acqua ne sono un buon esempio. Si immaginino due serie identiche di onde generate da sassi gettati in uno stagno. Esse si propagano sulla superficie l'una verso l'altra e l'altezza dell'acqua nel punto in cui si incontrano è la somma delle altezze delle singole onde. Quando la cresta di un'onda incontra la cresta di un'altra, l'onda risultante è più alta; quando, invece, una cresta incontra un ventre, le due onde si annullano.

È possibile sfruttare il fenomeno dell'interferenza per misurare la distanza relativa percorsa da due onde. Le onde si sommano se hanno percorso esattamente la stessa distanza, ma si annullano se i loro cammini differiscono di una semilunghezza d'onda. I rivelatori di onde gravitazionali a interferometria si basano appunto su questo effetto. Due onde luminose identiche vengono generate dividendo un fascio laser in due fasci perpendicolari per mezzo di uno specchio semiriflettente (divisore di fascio). I due fasci vengono poi diretti su masse sospese libere di rispondere a una deformazione gravitazionale. Specchi montati sulle masse riflettono i fasci in maniera tale che essi si riuniscano in corrispondenza del divisore di fascio. La luminosità del fascio di uscita risultante, che è proporzionale al quadrato dell'ampiezza della combinazione delle due onde, misura la differenza fra le distanze percorse per raggiungere ognuna delle masse. Se un'onda gravitazionale proveniente dall'alto colpisse l'apparecchiatura, la distanza percorsa da uno dei fasci verrebbe abbreviata e simultaneamente aumenterebbe quella percorsa dall'altro, variando, di conseguenza, la luminosità all'uscita. Misurando accuratamente tale luminosità, sarebbe possibile distinguere spostamenti, dovuti alle onde gravitazionali, molto minori di una lunghezza d'onda.

La natura discreta della luce limita però la precisione delle misure di luminosità. Il laser che illumina l'interferometro emette rapidamente fotoni; l'errore relativo nel conteggio dei fotoni è pari all'inverso della radice quadrata del numero rilevato durante la misurazione, perciò si può ridurre l'errore usando un laser più potente. Il minimo spostamento della durata di un millisecondo misurabile da un semplice interferometro illuminato da un laser della potenza di un watt è di 10^{-14} metri. Gli interferometri sono in grado di misurare spostamenti molto più piccoli delle dimensioni di un atomo, e perfino del nucleo di un atomo, perché i fasci di luce sono riflessi da migliaia di miliardi di atomi sulla superficie degli specchi.

La possibilità di aumentare la lunghezza di un interferometro è il fattore chiave per migliorarne la sensibilità. Gli attuali prototipi hanno lunghezze che vanno da 1,5 metri (al MIT) a 40 metri (al Caltech). Le masse di controllo degli interferometri di LIGO saranno lontane quattro chilometri, 100 volte di più che nel prototipo sperimentale del Caltech.

La sensibilità dei rivelatori può essere migliorata anche facendo riflettere molte volte la luce fra gli specchi dell'interferometro. Il miglioramento di sensibilità è direttamente proporzionale al numero delle riflessioni, purché la luce non indugi tanto nel percorso da far sì che gli effetti delle onde gravitazionali vengano mediati. Per «piegare» il cammino della luce si possono usare due tecniche diverse: una linea ottica di ritardo (come è stato proposto da Weiss) e una cavità di Fabry-Perot (come è stato proposto da Drever). In una linea ottica di ritardo la luce passa attraverso un forellino in uno specchio adiacente al divisore di fascio e viene riflessa ripetutamente prima di riemergere dal medesimo forellino. In una cavità di Fabry-Perot la luce passa attraverso uno specchio parzialmente trasmittente e «si accumula» in una cavità risonante coperta da uno specchio totalmente riflettente, per uscire poi di nuovo

Un interferometro, composto da un laser, un divisore di fascio, due specchi che definiscono altrettanti cammini ottici perpendicolari tra loro e un fotorivelatore, è in grado di rilevare onde gravitazionali. Sotto al rivelatore, è indicato il diagramma del campo elettrico in funzione del tempo per ognuno dei fasci luminosi. Se i due cammini ottici differiscono di un ottavo di lunghezza d'onda, i campi elettrici dei fasci si combinano con uno sfasamento di un quarto di ciclo e la luce che appare al forellino di uscita è metà dell'intensità luminosa totale (*in alto*). Se un'onda gravitazionale produce una variazione nei cammini ottici, la luce può interferire o distruttivamente, e l'intensità luminosa all'uscita sarà minore (*al centro*), o costruttivamente, dando origine a un'emissione luminosa più intensa (*in basso*).

La sensibilità di un interferometro a piccole variazioni nella posizione dei suoi specchi, prodotte da onde gravitazionali, può essere migliorata facendo riflettere più volte i fasci luminosi. Si possono impiegare due diverse configurazioni. In una linea ottica di ritardo (*in alto*), la luce entra attraverso un forellino in uno specchio, subisce una serie di riflessioni ed esce attraverso lo stesso forellino. Per semplicità, sono mostrati solo due passaggi; in realtà, le linee di ritardo usate nei rivelatori di onde gravitazionali producono decine di riflessioni (*al centro*). In una cavità di Fabry-Perot (*in basso*), la luce entra attraverso uno specchio parzialmente trasmittente, è riflessa da un secondo specchio e ritorna allo specchio di ingresso. Una piccola parte della luce incidente è trasmessa, mentre il resto è riflesso e torna in ciclo. Il numero di riflessioni dipende dalla riflettività dello specchio di ingresso; alcuni prototipi producono fino a 10 000 riflessioni.

attraversando lo specchio di ingresso.

Per cercare segnali vicini a una specifica frequenza, il «ritardo» della luce può essere regolato in modo che sia pari a un semiperiodo dell'onda gravitazionale. La sensibilità può essere aumentata sia alternando i cammini dei due fasci con frequenza uguale a quella cercata, sia riportando all'ingresso la luce che esce dall'interferometro, aumentando così, in pratica, la potenza del laser.

L'effetto della radiazione gravitazionale è «sentito» da tutte le masse dell'universo; le masse di controllo presenti nell'interferometro si distinguono solo per tre aspetti: sono libere di «rispondere» al passaggio di un'onda, sono contenute in uno strumento sensibile ai minuscoli spostamenti indotti dall'onda stessa e sono isolate da qualsiasi disturbo che possa coprire il segnale.

Tra i numerosi tipi di rumori che possono disturbare la risposta dei rivelatori alla radiazione gravitazionale vi sono i movimenti sismici, le vibrazioni termiche all'interno delle masse di controllo e gli effetti delle molecole residue di gas che colpiscono le masse stesse e penetrano nei fasci laser. Questi rumori di fondo sono in grado di produrre negli osservatori di onde gravitazionali spostamenti spuri molte volte maggiori di quelli prodotti dai segnali. I movimenti più cospicui del suolo, provocati da terremoti e dall'eventuale attività di persone o macchine nei pressi, sono infrequenti e comunque ben identificabili; maggiori problemi vengono posti invece dal costante fondo sismico debole, alimentato da venti e onde e dai moti all'interno della Terra.

L'intensità del fondo sismico diminuisce rapidamente a frequenze superiori a 10 hertz, ma continua a essere ben rilevabile anche a frequenze di centinaia di hertz. Per neutralizzare questo effetto indesiderato, le masse di controllo in un osservatorio di onde gravitazionali sono sospese a fili sottili, come i pesi di un pendolo. Eventuali movimenti rapidi al punto di sospensione vengono così molto attenuati prima di raggiungere la massa sottostante. A loro volta, i punti di sospensione sono isolati da una struttura di sostegno molleggiata, che attenua i sussulti dell'attività sismica funzionando in modo analogo alle sospensioni di un'automobile. Questa combinazione di fili e molle è in grado di annullare i rumori di origine sismica di frequenza superiore a 500 hertz, mentre per rilevare onde gravitazionali a frequenze più basse diventa inevitabile il ricorso a meccanismi di sospensione più complessi basati su sistemi a retroazione attiva.

Un'altra fonte di rumore sono le vibrazioni termiche all'interno delle masse di controllo. L'effetto di disturbo può essere attenuato montando masse compatte e di struttura semplice, che risuonino a frequenze molto superiori a quelle prevedibili per le onde gravitazionali.

La zona in cui si intersecano i tubi sotto vuoto, in quasi tutti gli interferometri, contiene gran parte degli strumenti ottici, che, se fossero riuniti in una struttura unica, presenterebbero un numero elevato di frequenze di risonanza contigue. Questa struttura centrale viene perciò solitamente suddivisa in componenti separati, ognuno con frequenze di risonanza tollerabili.

Il fatto che gli apparecchi siano contenuti in tubi sotto vuoto minimizza gli effetti delle molecole residue di gas che colpiscono le masse e penetrano nei fasci laser. È necessario però che il vuoto sia quasi perfetto, perché, se il fascio di luce incontra molecole residue di gas, la sua velocità non risulta costante.

Disturbi di minore entità derivano da variazioni locali del campo magnetico e gravitazionale. Ma è poco probabile che qualsiasi fonte di rumore, identificata o no, abbia effetti simultanei su rivelatori lontani migliaia di chilometri, come la coppia di rivelatori che compone LIGO. Le onde gravitazionali di origine cosmica investono uniformemente la Terra, perciò vengono rilevate quasi simultaneamente da tutti gli osservatori. È difficile immaginare che la «firma» di un impulso di radiazione gravitazionale, registrata contemporaneamente in due o più osservatori lontani, non venga individuata come tale.

Gli attuali interferometri da laboratorio non hanno una sensibilità sufficiente per rilevare onde gravitazionali provenienti da galassie lontane; d'altra parte, gli impulsi di origine galattica, che sono alla portata degli strumenti esistenti, si producono raramente, forse una volta ogni 30 anni. Con la sua lunghezza utile di quattro chilometri, LIGO sarà abbastanza sensibile da poter registrare impulsi che producano spostamenti di meno di 10^{-21} metri, un livello prevedibile per molti segnali di origine extragalattica. Inoltre, quasi tutte le fonti di rumore che disturbano le misurazioni, comprese l'attività sismica e l'agitazione termica, sono indipendenti dalla lunghezza del dispositivo, mentre l'intensità del segnale registrato aumenta con la lunghezza. Fonti di rumore come le fluttuazioni nel puntamento del laser e le deflessioni e le attenuazioni dovute alle molecole di gas, che diventano anch'esse maggiori con la lunghezza, sono comunque intrinsecamente piccole o facili da controllare.

Un solo rivelatore di onde gravitazionali ha una direzionalità molto bassa. Questo significa che interferometri posti in regioni molto distanti del globo possono registrare simultaneamente segnali provenienti da una vasta porzione di cielo, ma anche che occorrono almeno tre rivelatori per stabilire esattamente da quale direzione provengano gli impulsi, mentre per distinguere un segnale dal rumore di fondo ne bastano due. Un'informazione direzionale può essere ottenuta confrontando i tempi di arrivo del segnale ai vari rivelatori. Due grandi interferometri europei potrebbero entrare in servizio quasi contemporaneamente all'osservatorio statunitense. Uno di essi, con una lunghezza utile di un chilometro, aumentabile a tre chilometri, verrà costruito in Scozia dall'Università di Glasgow e dal Rutherford Appleton Laboratory. Il secondo, progettato in Baviera dagli Istituti Max Planck della Germania Occidentale, avrà percorsi ottici lunghi tre chilometri con un angolo di intersezione di 60 gradi; l'osservatorio potrà essere ampliato, assumendo una configurazione con tre interferometri disposti a triangolo equilatero. Un terzo grande interferometro, la cui entrata in servizio è prevista in tempi più lunghi, verrà costruito da un gruppo francese.

LIGO avrà tubi sotto vuoto abbastanza grandi (con un diametro di 1,2 metri) da contenere parecchi interferometri, di lunghezze diverse, ognuno con masse di controllo sospese separatamente. Tutti gli interferometri reagiranno allo stesso modo ai rumori, ma quelli di lunghezza doppia registreranno i segnali con intensità doppia; in questo modo sarà possibile distinguerli dagli eventi spuri.

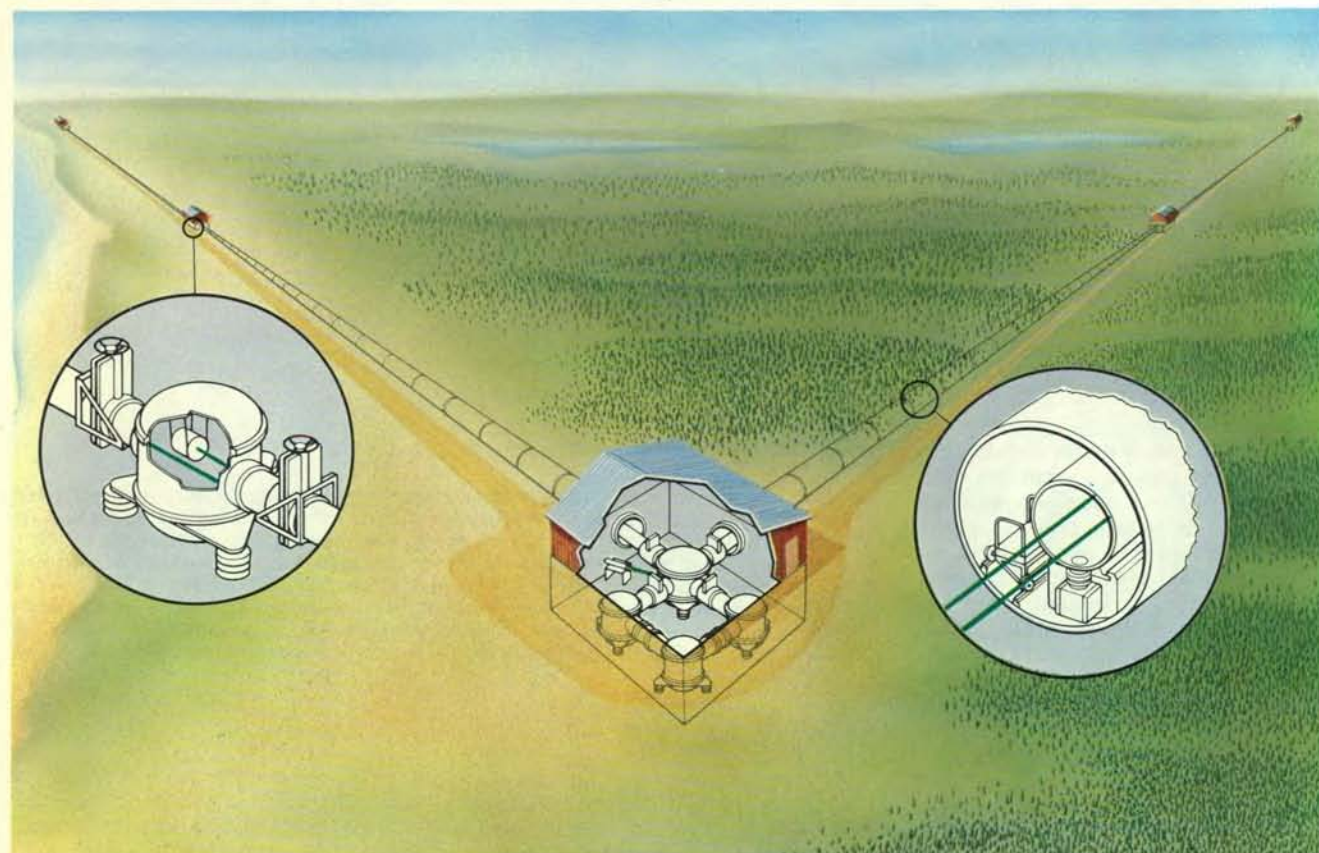
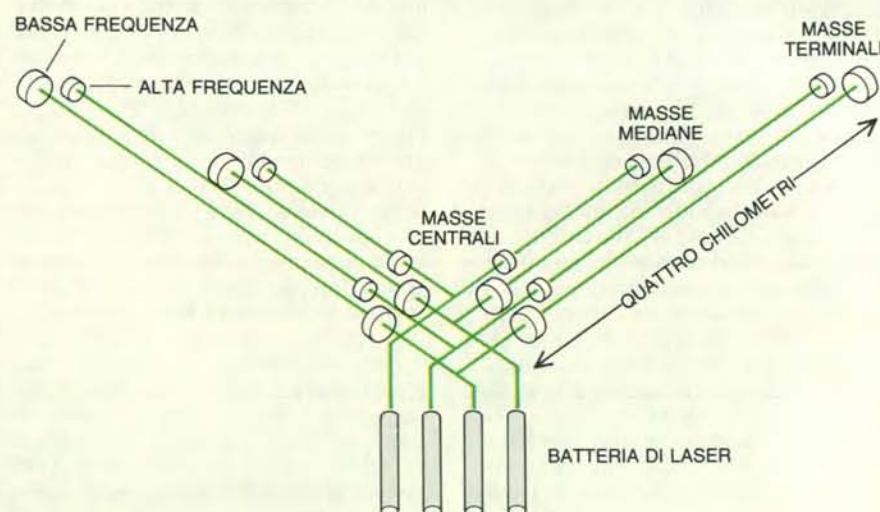
Poiché gran parte del rumore che disturba le misurazioni è prodotto dall'ambiente terrestre, sarebbe interessante installare osservatori di onde gravitazionali nello spazio. In effetti, le proposte in questo senso sono numerose. Un esperimento consiste nel misurare accuratamente la frequenza di un segnale di microonde riflesso verso la Terra da una sonda spaziale. Un eventuale impulso gravitazionale provocherebbe uno spostamento della frequenza delle microonde. Il metodo consentirebbe di rilevare onde gravitazionali con frequenze comprese fra una al minuto e una all'ora, molto inferiori all'intervallo di frequenze di un'antenna posta sulla Terra; la sensibilità degli strumenti attuali, tuttavia, non ha permesso di rilevare ancora alcun segnale.

Un altro possibile esperimento, più

Le masse di controllo in un rivelatore di onde gravitazionali a interferometria devono essere isolate da fonti di disturbo esterne, come i movimenti sismici. Per questa ragione, ogni massa, che può essere ridotta al solo specchio dell'interferometro, è sospesa a fili molto sottili, i cui capi opposti sono collegati a una struttura fatta di strati alterni di piombo e gomma. Per mantenere fissa l'orientazione angolare dello specchio, un fascio laser ausiliario è riflesso dalla superficie dello specchio su un fotorivelatore per il controllo della posizione. Il rivelatore invia segnali a trasduttori elettromagnetici che, agendo su una massa più piccola posta vicino al punto di sospensione, permettono di correggere le deviazioni angolari che interferiscono con le misurazioni.

elaborato, prevede la misurazione per mezzo del laser della separazione tra masse di controllo poste in orbita intorno al Sole; la distanza tra le masse potrebbe essere pari a un milione di chilometri. Tale sistema permetterebbe lo studio delle onde gravitazionali provenienti da sistemi binari classici, con lunghi periodi di rivoluzione. Un interferometro nello spazio, con una base molto più corta, potrebbe rilevare segnali nella banda dei chilohertz, ma le sue prestazioni non dovrebbero superare quelle dei rivelatori a terra, mentre la spesa prevedibile sarebbe molto maggiore.

Ogni nuova tecnologia introdotta nel campo astronomico, dai rivelatori radio, X e infrarossi ai telescopi ottici più moderni, ha aperto prospettive inedite nello studio del cosmo. Mentre gli astronomi del secolo scorso immaginavano l'universo come un aggregato tranquillo e immutabile di stelle, oggi sappiamo che esso nacque da una titanica esplosione e che si sta tuttora espandendo a velocità enorme. A loro volta, galassie e ammassi di galassie sono sede di fenomeni violenti, come l'espulsione di getti di gas lunghi migliaia di anni luce, la distruzione di stelle vecchie a causa di maree prodotte dalle compagne e la nascita di stelle all'interno delle nubi oscure.

LIGO e gli altri futuri osservatori di onde gravitazionali saranno in grado di fornirci per la prima volta un'immagine chiara di processi che possiamo solo intravvedere osservando nel campo delle radiazioni elettromagnetiche. Oltre a individuare eventi come il collasso del nucleo di una supernova, la fusione relativistica di stelle di neutroni e forse anche la formazione di buchi neri, potranno arricchire in maniera più determinante la nostra visione dell'universo con la scoperta di fenomeni del tutto nuovi.

Il grande interferometro, la cui entrata in servizio è prevista per gli anni novanta, farà parte di una coppia nota nel suo insieme come osservatorio a interferometria laser di onde gravitazionali, o LIGO. Uno degli interferometri verrà costruito nella California meridionale, l'altro nel Maine; i due strumenti verranno collegati elettronicamente in modo da poter funzionare come un singolo osservatorio. Ognuno degli strumenti avrà bracci di quattro chilometri di lunghezza. Uno dei progetti proposti per LIGO prevede l'installazione di quattro interferometri, ognuno dotato di un proprio dispositivo laser, riuniti in un tubo sotto vuoto (*a destra*). La possibilità di contenere in uno stesso osservatorio interferometri diversi, destinati a rilevare differenti segnali di origine cosmica, aumenterà la versatilità di LIGO. Ciascuno dei sistemi ottici verrà duplicato in una configurazione di lunghezza dimezzata (pari a due chilometri), per facilitare l'individuazione di segnali spuri.

L'anatomia della memoria

Studi sull'amnesia umana hanno dimostrato che le strutture profonde del cervello possono interagire con le vie della percezione degli strati cerebrali superficiali in modo da trasformare gli stimoli sensoriali in ricordi

di Mortimer Mishkin e Tim Appenzeller

Entro le pur limitate dimensioni del cervello umano esiste un sistema di memoria abbastanza potente da catturare l'immagine di un volto fin dal primo incontro, abbastanza capace da immagazzinare le esperienze di un'intera esistenza e così versatile che un luogo può richiamare associazioni di immagini, suoni, odori, sapori, sensazioni tattili ed emozioni. Come lavora questo sistema di memoria? Anche solo definire la memoria costituisce un problema; la riflessione suggerisce che esista una differenza fra il conoscere un volto o una poesia e l'acquisire un'abilità come lo scrivere a macchina. Inoltre, il substrato fisico della memoria, i cento miliardi o quasi di cellule nervose del cervello e l'intreccio delle loro connessioni, è enormemente complesso. Tuttavia, in questo articolo tenteremo di descrivere schematicamente come il cervello ricorda.

Il quadro cui siamo giunti è soprattutto anatomico. Nel corso di ricerche durate vent'anni abbiamo identificato le strutture nervose e le stazioni (vasti ammassi di cellule) che contribuiscono alla memoria, abbiamo tracciato le loro connessioni e cercato di determinare come interagiscono allorché un ricordo viene immagazzinato, richiamato o collegato ad altre esperienze. Altri ricercatori si sono dedicati a un'analisi più fine della memoria: in alcuni fra gli animali più semplici e nel tessuto nervoso isolato dagli animali superiori essi hanno messo in evidenza cambiamenti delle proprietà elettriche e chimiche di singoli neuroni come risultato di forme semplici di apprendimento. La complessità dell'oggetto della nostra ricerca, la memoria dell'uomo o, nella migliore approssimazione, delle scimmie del Vecchio Mondo, richiede un metodo di studio diverso, che evidenzi un'architettura su vasta scala. Infine, la memoria è certamente una successione di eventi molecolari. Ciò che dobbiamo definire è il territorio entro cui questi eventi hanno luogo.

Le conoscenze attuali provengono da osservazioni condotte su pazienti, i quali, in seguito a malattie, lesioni od operazioni chirurgiche riguardanti aree cerebrali specifiche, hanno perso una parte della loro capacità di apprendere e ricordare. Probabilmente il più famoso di questi casi è quello di un paziente affetto da amnesia noto come H. M.; studiato da Brenda Milner al Montreal Neurological Institute e da altri, H. M. ha fornito una grande quantità di informazioni sul quadro dei disturbi associati a un tipo specifico di danno cerebrale.

Il lavoro sperimentale, condotto per lo più sul macaco, ha unito ricerche anatomiche, fisiologiche e comportamentali. Sostanze traccianti che vengono trasportate lungo gli assoni, i sottili prolungamenti attraverso i quali i neuroni inviano segnali, hanno rivelato il circuito nervoso che potrebbe essere responsabile del ruolo svolto da strutture specifiche nella memorizzazione. Misurazioni dell'attività elettrica dei neuroni, o l'assunzione da parte dei neuroni stessi di glucosio marcato hanno permesso di distinguere le regioni del cervello attivate nel corso di esercizi che richiedono un apprendimento. Il gruppo finale di esperimenti, volti a stabilire l'importanza funzionale di strutture identificate con altri mezzi, ha richiesto sia interventi chirurgici o somministrazione di farmaci, sia test psicologici. Nel cervello di un animale sottoposto a sperimentazione, possono venire distrutti o bloccati dai farmaci interi gruppi di neuroni oppure possono venire interrotte le vie che li collegano. L'animale può essere poi sottoposto a test comportamentali per separare le varie componenti della memoria e stabilire quale di esse sia danneggiata.

Il nostro metodo per affrontare lo studio della memoria umana è indiretto e presenta inevitabili problemi. Il cervello di un macaco è circa un quarto di quello di uno scimpanzé, il parente più prossimo dell'uomo, e il cervello di uno scimpanzé, a sua volta, è solo un quarto circa di quello umano. Con l'aumentare delle dimensioni è aumentata anche la complessità. Le strutture che studiamo nel macaco hanno tutte un corrispettivo nel cervello umano, ma le loro funzioni possono essersi differenziate nel corso dell'evoluzione. Soprattutto la capacità di parlare, tipica dell'essere umano, e le specializzazioni cerebrali che a essa si accompagnano pongono dei limiti allo studio comparativo. Tuttavia i sistemi nervosi di base sono probabilmente gli stessi nelle scimmie e nell'uomo e i nostri risultati confermano ciò che già sappiamo sulla perdita della memoria nell'uomo.

Il sistema visivo

Molto spesso i ricordi hanno origine da impressioni sensoriali. Prima di chiedersi come il cervello possa immagazzinare un'esperienza sensoriale sotto forma di ricordo, bisognerebbe, anzitutto, conoscere come elabora questo tipo di informazione. Uno studio delle vie nervose responsabili della percezione visiva ha costituito infatti il punto di partenza della nostra indagine sulla memoria.

La corteccia striata, o corteccia visiva primaria, è un'area situata sulla superficie posteriore del cervello che riceve le informazioni riguardanti il mondo visivo, provenienti dalla retina, attraverso il nervo ottico e una stazione intermedia situata in profondità nel cervello (il corpo genicolato laterale). La corteccia striata è in grado di costruire una mappa sistematica del campo visivo: ogni porzione del campo attiva un gruppo distinto di neuroni. Tuttavia il sistema visivo non termina nella corteccia striata. Fin dagli anni cinquanta fu chiaro che il lobo temporale, una parte di ciascun emisfero cerebrale che si trova dietro l'orecchio e la tempia, ha un ruolo nella visione.

Alcuni ricercatori, fra cui uno di noi (Mishkin), scoprirono, negli anni sessanta, che l'area visiva del lobo temporale è in effetti il proseguimento di una via che ha inizio nella corteccia striata e si estende in avanti attraverso il tessuto

Oggetti chiaramente distinguibili servono per studiare la memoria delle scimmie. Uno degli autori (Mishkin) e collaboratori hanno cercato di identificare le strutture e le vie cerebrali che consentono a una scimmia di riconoscere, in una coppia, un oggetto visto una sola volta e per poco tempo. Utilizzando diversi oggetti i ricercatori si assicurano che l'animale impari a riconoscere un esemplare del tutto nuovo a ogni prova.

corticale nella corteccia temporale inferiore, sulla superficie inferiore del lobo temporale. Studi di neuroanatomia hanno dimostrato che numerose stazioni corticali vengono collegate in varie sequenze, lungo questa via.

Ricercatori di vari laboratori hanno studiato il ruolo esercitato nella percezione visiva da particolari stazioni ledendo chirurgicamente la via ottica in scimmie, sottoponendole poi a esercizi visivi e registrando l'attività elettrica di ciascuna stazione in animali esposti a stimoli visivi differenti. In un importante esperimento, Charles G. Gross, della Princeton University, e collaboratori hanno registrato le risposte di neuroni situati nella corteccia temporale inferiore a stimoli - forme di piccole dimensioni - che venivano mostrati alle scimmie.

Sebbene singoli neuroni situati nella corteccia striata - all'inizio della via visiva - rispondano più energicamente a uno stimolo semplice, come una barretta con orientazione e localizzazione specifiche nel campo visivo (fenomeno scoperto da David H. Hubel e Torsten N. Wiesel), i neuroni inferotemporali studiati da Gross e collaboratori rispondevano a forme più complesse poste lateralmente entro un'area di venti-trenta gradi. Alcuni neuroni rispondevano anche a una forma particolare posta in qualunque punto del campo visivo. I risultati facevano pensare che ciascun neurone inferotemporale ricevesse da vaste zone del campo visivo dati che riguardavano spesso l'intera gamma di proprietà che costituisce uno stimolo visivo.

Questi e altri risultati ci hanno portato a supporre che l'informazione visiva venga gradualmente elaborata lungo la via nervosa. I neuroni di questa via presentano «finestre» affacciate sul mondo esterno che divengono progressivamente più ampie, sia per quanto riguarda la loro estensione nello spazio sia per quanto riguarda la complessità dell'informazione che ricevono, in corrispondenza di stazioni successive. Le cellule rispondono progressivamente a varie proprietà fisiche di un oggetto - grandezza, forma, colore, struttura - finché, nelle stazioni terminali della corteccia temporale inferiore, riescono a sintetizzare una rappresentazione completa dell'oggetto.

Dall'esperienza alla memoria

Lungo la via visiva, quindi, il cervello integra i dati sensoriali in un'esperienza percettiva. Sembra che i dati relativi ad altri sensi vengano elaborati in maniera simile; nel laboratorio di uno di noi (Mishkin), al National Institute of Mental Health, David P. Friedman, Elisabeth A. Murray, Timothy P. Pons e Richard J. Schneider hanno identificato recentemente un'estesa via di elaborazione delle sensazioni tattili. Nella prima stazione singoli neuroni rispondono a punti localizzati sulla superficie corporea, mentre i neuroni della stazione terminale rispondono a stimoli provenienti da aree estese forse a tutta l'esperienza tattile.

Che cosa deve accadere perché percezioni così integrate vengano immagazzinate come ricordi? Il problema è ancora oggetto di ricerca, ma si possono già identificare le vie che hanno una funzione determinante al riguardo. Esse sono legate a due strutture situate sulla superficie interna del lobo temporale, in entrambi gli emisferi cerebrali: l'ippocampo, dalla parola greca che significa cavalluccio marino e l'amigdala, dalla parola greca mandorla.

Studi condotti su soggetti umani avevano già suggerito che l'ippocampo giocasse un ruolo determinante nei processi di memorizzazione. Dagli anni cinquanta, la rimozione chirurgica di parte del lobo temporale costituisce l'estremo tentativo di cura per pazienti affetti da crisi epilettiche gravi che si originano in quella zona del cervello. Inizialmente, alcuni pazienti manifestavano un'amnesia molto marcata dopo l'operazione (questa fu la causa dell'handicap di H. M.). Due caratteristiche della perdita di memoria erano degne di nota. Per prima cosa, l'amnesia era globale, comprendendo ricordi di esperienze che includevano tutte le modalità sensoriali, in secondo luogo era di tipo anterogrado, poiché, anche se i pazienti mantenevano i ricordi accumulati qualche tempo prima dell'operazione, non potevano ricordare alcunché di nuovo. In tutti i casi in cui si verificava questa perdita di memoria l'intervento aveva danneggiato l'ippocampo.

Si vide, tuttavia che era impossibile riprodurre un'analoga perdita di memoria nell'animale rimuovendo soltanto l'ippocampo. In una serie di esperimenti riguardanti, invece, l'amigdala, trovammo che questa struttura gioca un ruolo importante come quello dell'ippocampo nella memorizzazione. Danneggiando contemporaneamente le due strutture, in entrambi gli emisferi della scimmia, ottenemmo un animale che presentava amnesia anterograda globale.

Nella sezione del cervello umano vengono messe in evidenza le zone del cervello che, se danneggiate da malattie o da altri eventi, possono provocare una perdita di memoria (*in alto*). Le aree in colore indicano le strutture corrispondenti nel cervello del macaco (*in basso*). Mentre il cervello umano è ridotto di un terzo, quello del macaco è rappresentato in grandezza naturale.

Il sistema visivo elabora l'informazione lungo due vie che si trovano nella corteccia cerebrale. L'elaborazione iniziale dell'informazione (che arriva dalla retina attraverso il corpo genicolato laterale) ha luogo nel punto di partenza di queste vie, la corteccia striata. Qui singoli neuroni rispondono a elementi semplici, che occupano uno spazio limitato nel campo visivo, come contorni o macchie di colore. Lungo la via inferiore (che in effetti è formata da numerosi canali divergenti e di nuovo convergenti) i neuroni analizzano le proprietà macroscopiche di un oggetto, come, per esempio, la sua forma globale o il colore. Nel tratto finale di questa via, che corrisponde alla corteccia temporale inferiore, i singoli neuroni sono sensibili a un gran numero di proprietà e a un vasto campo visivo, il che suggerisce che proprio lì debba convergere l'informazione, elaborata nella sua completezza, concernente un oggetto. Lungo la via corticale superiore, che non è stata studiata così dettagliatamente, vengono analizzati i rapporti spaziali di una scena. La percezione della posizione di un oggetto rispetto ad altri punti di riferimento del campo visivo si formerebbe, per esempio, nella stazione terminale della via «spaziale», che è localizzata nella corteccia parietale posteriore.

Inizialmente, il nostro scopo era quello di capire perché un danno chirurgico alla corteccia temporale inferiore nella scimmia rendesse l'animale incapace di un tipo particolare di apprendimento visivo: la scelta di un oggetto o di una forma che erano stati costantemente associati con un premio in cibo rispetto ad altri non associati ad alcuna ricompensa. Noi pensavamo che la difficoltà fosse la conseguenza di un danno alla percezione visiva, che rifletteva un danno alla stazione di elaborazione più elevata del sistema visivo, ma era anche possibile che il danno riflettesse, invece, l'incapacità di associare uno stimolo con un premio.

Un modo per escludere questa possibilità - e confermare quindi che la difficoltà delle scimmie che presentavano lesioni nella regione temporale inferiore era di natura visiva - poteva essere quello di dimostrare che una struttura diversa, non danneggiata nelle scimmie sottoposte all'esperimento, era responsabile del collegamento tra premio e stimolo visivo. Con Barry Jones, allora alla McMaster University, nell'Ontario, uno di noi (Mishkin) cercò strutture che presentassero connessioni anatomiche con il sistema visivo e studiò le conseguenze della distruzione chirurgica di quelle strutture sulla capacità delle scimmie di scegliere un dato oggetto. Due delle strutture studiate furono l'amigdala e l'ippocampo, che presentano estese connessioni (indirette nel caso dell'ippocampo) con la corteccia temporale inferiore.

Il basso livello di prestazioni conseguente alla rimozione bilaterale dell'amigdala suggerì che essa fosse la struttura principale responsabile dell'associazione positiva - l'attesa del cibo - con uno stimolo elaborato dal sistema visivo. Prima di analizzare più in profondità l'interazione fra il sistema visivo e l'amigdala, insieme a Brenda J. Spiegler, dell'Università del Maryland a College Park, uno di noi (Mishkin) cercò di aumentare il grado di menomazione, estendendo chirurgicamente l'area lesa in modo da comprendere sia l'amigdala, sia l'ippocampo, la cui rimozione non aveva avuto, di per se stessa, alcun effetto.

Gli animali cui era stata rimossa soltanto l'amigdala apprendevano lentamente l'associazione fra stimolo e premio, ma erano ancora capaci di farlo. Fummo perciò sorpresi di scoprire che la rimozione combinata di amigdala e ippocampo cancellava del tutto nelle scimmie la capacità di compiere l'esercizio. Questo completo deterioramento delle prestazioni ci portò a chiederci se non si trattasse di un deficit che andava al di là dell'incapacità di associare un oggetto familiare con un premio. Il fatto che le scimmie non fossero più in grado di scegliere l'oggetto corretto era forse dovuto alla loro incapacità di ricordare quell'oggetto da una prova all'altra? Distruggendo l'ippocampo insieme all'amigdala era forse stata creata un'amnesia visiva?

Insieme a Jean Delacour, dell'Università di Parigi, avevamo da poco messo a punto un test la cui specificità nei confronti della memoria visiva era autonoma rispetto alla capacità di collegare un oggetto con un premio. In questo test di memoria, detto discriminazione ritardata non familiare fra due oggetti, all'animale viene presentato un dato oggetto, sotto cui trova un premio costituito da una nocciolina o da un pezzettino di banana. Poi all'animale vengono mostrati due oggetti, uno dei quali è quello che ha visto precedentemente, mentre l'altro non gli è familiare. Questa volta il cibo è nascosto sotto il nuovo oggetto; la scimmia viene perciò premiata allorché riconosce ed evita l'oggetto familiare in favore di quello nuovo. In ogni prova vengono usati due nuovi oggetti.

Si tratta di un esercizio che le scimmie

apprendono velocemente. David Gaffan, dell'Università di Oxford, ha utilizzato questo test per studiare la memoria visiva della scimmia, aumentando l'intervallo fra la presentazione iniziale e la scelta o presentando all'animale non un solo oggetto, ma tutta una serie di oggetti da ricordare, cui faceva seguito una serie di scelte. Poiché il cibo è sempre associato con un oggetto nuovo, la capacità di collegare un oggetto specifico con un premio non è significativa. Il premio rappresenta solamente un incentivo; il test misura in modo specifico la memoria di riconoscimento. Scimmie normali ottengono quasi sempre punteggi elevati.

Avevamo già applicato questo test di riconoscimento visivo a scimmie la cui corteccia temporale inferiore era stata rimossa e avevamo trovato che il test registrava la stessa difficoltà nel percepire o identificare oggetti familiari che avevamo rilevato nei test precedenti. Ripetemmo la prova su animali il cui sistema visivo era intatto, ma ai quali erano stati rimossi bilateralmente l'ippocampo e l'amigdala. Quando l'intervallo fra la presentazione dell'oggetto e la scelta era breve, gli animali riuscivano a eseguire l'esercizio, il che faceva pensare che non presentassero difetti nella percezione visiva. Quando aumentammo l'intervallo a un minuto o due, invece, i loro punteggi caddero quasi al livello di scelta casuale. Avevamo quindi provocato una vera e propria perdita di memoria.

Per di più, l'amnesia risultante da lesioni combinate di ippocampo e amigdala non era limitata agli stimoli visivi: sembrava essere globale. Nel laboratorio di uno di noi (Mishkin), Murray rilevò un peggioramento simile in scimmie sottoposte a un test basato sulla discriminazione tattile di oggetti diversi, cioè sulla capacità di riconoscere oggetti mediante il tatto. Si era pensato che, negli esseri umani, questa amnesia globale fosse la conseguenza di un danno al solo ippocampo e, infatti, recenti osservazioni effettuate da Larry R. Squire e collaboratori, dell'Università della California a San Diego, durante l'autopsia di un paziente affetto da amnesia, hanno mostrato l'esistenza di lesioni limitate all'ippocampo. Molti altri pazienti presentano, tuttavia, lesioni più estese e un danno all'amigdala può contribuire al deterioramento delle loro prestazioni. Uno studio condotto tempo fa su pazienti affetti da amnesia aveva suggerito che la gravità della loro perdita di memoria poteva variare in proporzione al danno congiunto di amigdala e ippocampo.

Continuando a esaminare, con Richard C. Saunders, il contributo apportato da ciascuna struttura alla memoria nella scimmia, appurammo l'esistenza di un effetto graduale. L'amigdala e l'ippocampo esercitano lo stesso ruolo nella mediazione della memoria di riconoscimento visivo; la rimozione di una delle due strutture ha di per sé un effetto trascurabile sulla capacità di riconoscimento, probabilmente perché una struttura può sostituirsi all'altra. La rimozione di una delle due strutture da entrambi gli emisferi, e dell'altra da un solo emisfero, provoca un deterioramento molto più evidente. La rimozione bilaterale di amigdala e ippocampo, infine, fa sì che l'animale ottenga, nel test della discriminazione ritardata non familiare, punteggi solo di poco superiori rispetto a quelli che otterrebbe con una scelta casuale.

Altre stazioni

Il danno all'amigdala e all'ippocampo, due componenti importanti del sistema limbico, non è la sola neuropatologia che può dare luogo ad amnesia globale. In altri pazienti, il danno è localizzato nel diencefalo, un insieme di nuclei situato più o meno al centro del cervello, a sua volta organizzato in due strutture, il talamo e l'ipotalamo. Parti del diencefalo situate presso la linea mediana del cervello degenerano nella sindrome di Korsakoff, un'amnesia globale che si osserva in alcuni alcolizzati cronici; danni diencefalici provocati da ictus, lesioni, infezioni e tumori possono dar luogo allo stesso tipo di sindrome amnesica. Il fatto che i nuclei diencefalici siano coinvolti nel processo di memorizzazione è confermato sia da osservazioni cliniche sia da riscontri anatomici che mostrano come il diencefalo riceva fibre provenienti dall'ippocampo e dall'amigdala.

Per mettere alla prova l'ipotesi che il diencefalo interagisca con le strutture limbiche in una sorta di circuito della memoria, applicammo nuovamente la nostra strategia sperimentale chirurgica seguita da prove comportamentali. Con John P. Aggleton, dell'Università di Durham, uno di noi (Mishkin) distrusse, prima congiuntamente e poi separatamente, le regioni del diencefalo alle quali afferiscono ippocampo e amigdala. Prove di riconoscimento visivo misero in evidenza lo stesso tipo di danno alla memoria che era stato osservato dopo la rimozione di ippocampo e amigdala. Danni congiunti alle aree di proiezione diencefalica di entrambe le strutture disturbarono molto la memoria di riconoscimento delle scimmie, mentre danni limitati a una sola di queste aree avevano uno scarso effetto. Fu chiaro che avevamo identificato due distinti circuiti della memoria, ognuno dei quali era sufficiente per il riconoscimento visivo.

Ulteriori ricerche hanno confermato che il diencefalo e le strutture limbiche fanno parte di un circuito anziché dare un contributo individuale alla memoria; nel laboratorio di uno di noi (Mishkin), Jocelyne H. Bachevalier e John K. Parkinson produssero lo stesso deficit di memoria osservato in precedenza interrompendo le connessioni di queste strutture.

Precedenti ricerche neuroanatomiche facevano pensare che non fossimo ancora riusciti a seguire i circuiti per tutta la loro estensione. Nuclei talamici comunicanti con le strutture limbiche inviano fibre anche alla corteccia prefrontale ventromediale, una zona della corteccia situata sotto la parte anteriore del cervello. Anche a questo livello la Bachevalier osservò che lesioni chirurgiche determinavano una perdita notevole della memoria di riconoscimento.

In tal modo, la stazione finale del sistema visivo, come di altri sistemi sensoriali, risulta collegata a due circuiti di memoria paralleli che comprendono almeno le strutture limbiche del lobo temporale, le parti mediali del diencefalo e la corteccia prefrontale ventromediale. Come cooperano queste strutture nel processo di memorizzazione? Il problema è complicato dal fatto che i ricordi non sono probabilmente immagazzinati solo, o principalmente, nei circuiti stessi. Osservazioni cliniche hanno dimostrato che un danno al sistema limbico umano lascia intatti e accessibili i ricordi passati, il che fa ritenere che essi possano venire immagazzinati anche in corrispondenza di una stazione precedente i circuiti nervosi che abbiamo identificato. I depositi più probabili dei ricordi, sono, infatti, le stesse aree corticali dove prendono forma le impressioni sensoriali.

I circuiti subcorticali devono quindi essere coinvolti in una sorta di processo di retroazione con la corteccia. Dopo che uno stimolo sensoriale elaborato attiva l'amigdala e l'ippocampo, i circuiti mnestici devono, a loro volta, agire sull'area sensoriale. Questo processo probabilmente rafforza e quindi ha forse l'effetto di immagazzinare la rappresentazione nervosa dell'evento sensoriale che ha appena avuto luogo. La rappresentazione stessa probabilmente assume la forma di un insieme di molti neuroni, connessi fra loro in maniera particolare. Come risultato del processo di retroazione da parte dei circuiti mnestici, le sinapsi (le connessioni fra le cellule nervose) di questo gruppo nervoso potrebbero andare incontro a cambiamenti che conserverebbero lo schema delle connessioni, trasformando la percezione in un ricordo duraturo. Il riconoscimento avrebbe luogo più tardi, allorché l'insieme di neuroni viene riattivato dallo stesso evento sensoriale che lo ha originato.

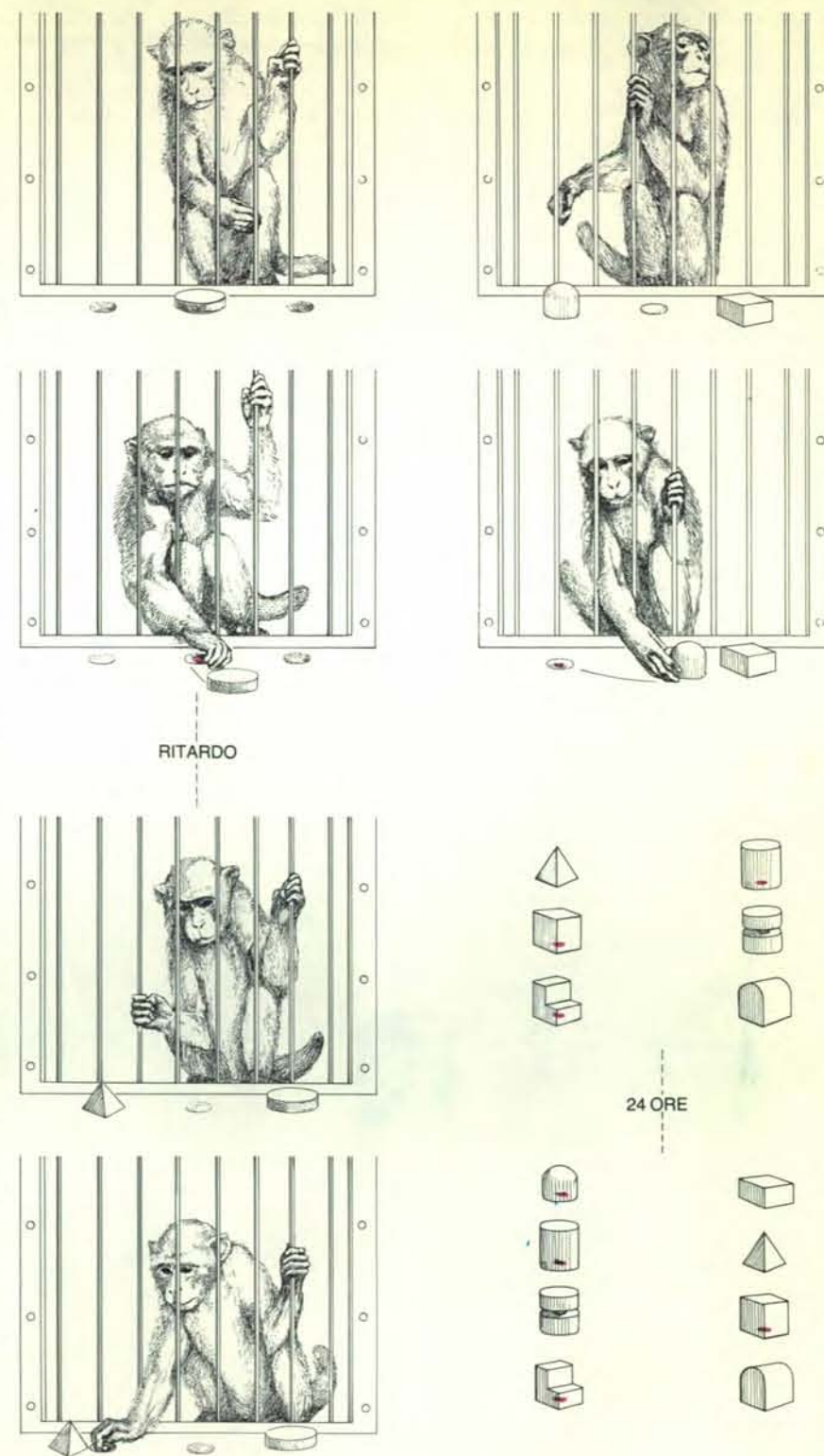
Non è noto come ciascuna struttura dei circuiti mnestici possa contribuire al processo descritto. Esistono tuttavia alcune indicazioni sulla natura del processo nel suo insieme. Una di queste riguarda un'altra struttura che, in base alle nostre ricerche, è risultata implicata nella memoria di riconoscimento. Si tratta del sistema colinergico situato nella regione proencefalica inferiore, un insieme di neuroni che fornisce alla corteccia e al sistema limbico il principale apporto di acetilcolina, un neurotrasmettitore.

Sembra che l'acetilcolina svolga un ruolo determinante nella memorizzazione. Livelli di acetilcolina inferiori al normale si registrano nel morbo di Alzheimer che ha tra i suoi sintomi più caratteristici la perdita della memoria. Nel laboratorio di uno di noi (Mishkin), Thomas G. Aigner ha trovato che le prestazioni delle scimmie migliorano se esse vengono trattate con fisostigmina, una sostanza che potenzia l'azione dell'acetilcolina. Quando, invece, viene loro somministrata scopolamina, una sostanza che blocca l'azione dell'acetilcolina, le loro prestazioni risultano disturbate. Recentemente, un gruppo di ricercatori diretto da Donald L. Price e Mahlon R. DeLong, della Johns Hopkins University School of Medicine, in collaborazione con Aigner e uno di noi (Mishkin), ha dimostrato che una lesione alla regione proencefalica inferiore danneggia la memoria di riconoscimento nelle scimmie, sebbene l'effetto finora osservato non sia così grave e duraturo come quello alle altre strutture che abbiamo studiato.

Certamente anche in questa regione è presente un circuito che contribuisce alla formazione dei ricordi. Per esempio, l'ippocampo e l'amigdala inviano un gran numero di proiezioni a questa zona del cervello, che, a sua volta, rimanda fibre contenenti acetilcolina non solo alle strutture del sistema limbico, ma anche alla corteccia. In una rappresentazione verosimile di come si formano i ricordi, l'attivazione dei circuiti mnestici sottocorticali da parte di uno stimolo sensoriale provocherebbe la liberazione di acetilcolina dalla regione proencefalica inferiore verso l'area sensoriale. L'acetilcolina (e forse anche altri neurotrasmettitori) darebbe inizio a una serie di eventi cellulari che modificherebbero le sinapsi del tessuto nervoso, rafforzerebbero le connessioni nervose e trasformerebbero la percezione sensoriale in una traccia mnestica stabile.

Recenti ricerche biochimiche fanno ritenere che nelle stazioni terminali del sistema visivo, dove è probabile venga immagazzinata la memoria, operi un meccanismo di modificazione sinaptica. Aryeh Routtenberg, della Northwestern University, ha proposto che l'aggiunta di un gruppo fosfato alla proteina cerebrale nota come *F1* da parte di un enzima, la proteinchinasi *C*, sia alla base dei cambiamenti sinaptici osservati in seguito a stimolazione elettrica ripetuta di alcuni neuroni. Per misurare l'attività del meccanismo di fosforilazione nel sistema visivo della scimmia, Routtenberg e il suo allievo Robert B. Nelson hanno addizionato fosforo radioattivo a tessuto sezionato dalle aree visive. Le quantità più elevate di tracciante risultarono incorporate nella proteina *F1* del tessuto ottenuto dalle stazioni finali. Questa osservazione può indicare che il tessuto è particolarmente suscettibile a cambiamenti sinaptici che potrebbero immagazzinare i ricordi.

Tipi di memoria

L'architettura cerebrale descritta è stata messa in evidenza per il suo contri-

Due prove possono permettere di valutare capacità di apprendimento fondamentalmente diverse. Nel test noto come discriminazione ritardata non familiare (*a sinistra*) viene presentato alla scimmia un oggetto non familiare, che deve prendere per trovare un premio. Dopo qualche tempo, posto davanti allo stesso oggetto insieme a uno nuovo, l'animale deve riconoscere l'oggetto che ha visto prima e spostare quello nuovo per ottenere il premio. Quando la scimmia ha imparato questo compito, l'esercizio può venire reso più difficile (aumentando, per esempio, l'intervallo di tempo) per mettere alla prova la capacità dell'animale di ricordare un oggetto visto una volta soltanto. Un secondo test, noto come discriminazione fra due oggetti familiari (*a destra*), impiega una successione di 20 coppie di oggetti. Un oggetto di ciascuna coppia nasconde il premio. La stessa sequenza è mostrata a una scimmia più volte, a intervalli di 24 ore, finché essa impara a scegliere l'oggetto giusto di ogni coppia. Uno degli autori (Mishkin) e collaboratori hanno dimostrato che, anche danneggiando le strutture cerebrali fondamentali per eseguire bene il primo esercizio, non viene menomata la capacità della scimmia di apprendere il secondo esercizio.

buto a un tipo di memoria: la memoria di riconoscimento, l'importante facoltà per cui una scimmia, dopo avere visto o toccato un oggetto una sola volta, può riconoscerlo parecchi minuti più tardi e trascurarlo in favore di uno nuovo. Naturalmente esistono forme più complesse di memoria, alcune delle quali possono anche venire studiate nelle scimmie. Il loro esame ha rivelato interessanti specializzazioni entro le vie nervose.

Allorché un oggetto viene percepito, per esempio, si impara non solo a distinguerne le caratteristiche particolari, ma anche a localizzarlo rispetto ad altri oggetti o punti di riferimento. Il ricordo di una scultura è, intuitivamente, un compito diverso dal ricordo della sua collocazione in una sala da esposizione. Il problema è diverso anche dal punto di vista anatomico. La capacità di osservare le relazioni spaziali dipende da una parte del sistema visivo differente da quella responsabile della percezione delle caratteristiche peculiari di un oggetto.

Nel 1973, Walter Pohl che allora lavorava nel laboratorio di uno di noi (Mishkin), confermò che il tessuto della corteccia parietale ha un ruolo nella visione. Egli dimostrò che la rimozione di questo tessuto causa un disturbo visivo che è molto differente da quello provocato da un danno alla corteccia temporale inferiore. A differenza delle scimmie con lesioni inferotemporali, gli animali con danni alla corteccia parietale sono ancora in grado di distinguere gli oggetti, ma non di percepire le relazioni spaziali.

Per i suoi esperimenti, Pohl utilizzò scimmie alle quali venivano presentate due vaschette coperte contenenti cibo. Fra le vaschette era posto un oggetto cilindrico, che si trovava più vicino a una delle due, in una posizione che veniva variata da una prova all'altra. Le scimmie venivano premiate quando scoprivano la vaschetta più vicina al cilindro, che conteneva una nocciolina, mentre l'altra era vuota. L'esercizio è relativamente facile per animali che presentano lesioni inferotemporali. Gli animali con danni alla corteccia parietale posteriore, invece, mostravano una difficoltà notevole a imparare qual era la vaschetta giusta.

Ulteriori ricerche, fra cui uno studio biochimico compiuto da Kathleen A. Macko, Charlene D. Jarvis e uno di noi (Mishkin), insieme a un gruppo diretto da Charles Kennedy e Louis Sokoloff, anch'essi del National Institute of Mental Health, confermarono che la corteccia parietale posteriore fa parte del sistema visivo. Iniettammo un analogo marcato del glucosio e ne studiammo l'assunzione nel corso di esperimenti collegati con la visione; osservammo così che i processi visivi coinvolgevano non solo la corteccia temporale inferiore, ma anche il tessuto parietale posteriore.

Inoltre, Leslie G. Ungerleider ha recentemente messo in evidenza che vi è un'altra via anatomica, oltre a quella nota per essere implicata nell'elaborazione delle caratteristiche visive di un oggetto, che ha origine dalla corteccia striata (la stazione visiva primaria situata nella parte posteriore del cervello). Invece di procedere nella corteccia parietale inferiore, questa seconda via sale attraverso una serie di stazioni fino alla stazione terminale situata nella corteccia parietale posteriore. Le relazioni spaziali vengono probabilmente analizzate lungo questa via; dalla stazione terminale le percezioni spaziali elaborate attivano presumibilmente il sistema di memorizzazione sottocorticale.

I ruoli dei due circuiti descritti possono differire nel caso dell'apprendimento spaziale. Sebbene l'ippocampo e l'amigdala siano in grado di sostituirsi l'uno all'altra quando si impara a riconoscere un oggetto, la prima struttura sembra essere particolarmente importante nell'apprendimento dei rapporti spaziali. Parkinson, dopo essersi interessato agli studi di altri ricercatori, i cui risultati avevano sottolineato l'importanza dell'ippocampo nell'apprendimento spaziale dei roditori, si dedicò allo studio di questa struttura cerebrale nelle scimmie.

Egli allenò scimmie normali a eseguire una prova che consentiva di valutare la loro memoria nella localizzazione di oggetti. A ogni prova venivano mostrati agli animali due oggetti completamente nuovi in posizioni specifiche su un vassoio; veniva poi presentato loro uno degli oggetti già visti nella sua posizione iniziale e un suo duplicato esatto, situato o dove era stato posto il secondo oggetto, oppure in una terza posizione. Gli animali venivano ricompensati quando sceglievano l'oggetto mostrato per primo nella sua posizione iniziale.

Dopo la rimozione bilaterale dell'amigdala, gli animali riapprendevano velocemente l'esercizio e lo eseguivano correttamente; se, invece, era l'ippocampo a venire asportato, le scimmie perdevano la capacità di compiere l'esercizio. Mary Lou Smith del Montreal Neurological Institute, lavorando con Milner, ha di recente riferito che un fatto simile si verifica anche in pazienti affetti da amnesia: vi è cioè una correlazione fra l'estensione del danno all'ippocampo e l'entità del disturbo che riguarda il ricordo della localizzazione di oggetti.

Il sistema della memoria è stato identificato sottoponendo le scimmie a un test che misura la memoria di riconoscimento visivo (*si veda l'illustrazione di pagina 39 a sinistra*), dopo aver provocato danni chirurgici a strutture o vie cerebrali specifiche. Il disegno in alto indica le strutture che hanno un ruolo determinante in questo tipo di memoria; il grafico al centro mostra i punteggi medi raggiunti dalle scimmie nel test dopo lesioni chirurgiche a elementi specifici del sistema; per confronto, esso mostra anche le prestazioni quasi perfette di scimmie normali. Il grafico in basso (basato su numerose prove, fra cui studi neuroanatomici del circuito che collega le strutture) mostra come le strutture possano interagire nella formazione di un ricordo. Una sensazione percepita nella stazione terminale del sistema corticale attiva due circuiti paralleli, uno dei quali ha origine nell'amigdala e l'altro nell'ippocampo; entrambi comprendono parti del diencefalo e della corteccia prefrontale. Ciascuna struttura invia, a sua volta, segnali alla regione proencefalica inferiore. Attraverso le sue numerose connessioni con la corteccia, tale regione potrebbe chiudere il circuito, dando luogo a cambiamenti dei neuroni dell'area sensoriale; in tal modo, la percezione potrebbe essere immagazzinata a questo livello come ricordo.

L'acquisizione di abitudini, ossia lo sviluppo di collegamenti automatici fra uno stimolo e una risposta, può dipendere dal corpo striato. Questa struttura riceve numerose proiezioni dai sistemi sensoriali corticali (nel disegno è evidenziato un ramo del sistema visivo) e, a sua volta, invia fibre alle strutture cerebrali che comunicano con la corteccia premotoria e la corteccia motoria supplementare. Attraverso le osservazioni neuroanatomiche, è stata individuata una via relativamente diretta che consente a uno stimolo giunto in un'area sensoriale di produrre una risposta motoria. Un danno al corpo striato, o alle connessioni fra questo e la corteccia, abbassa il livello delle prestazioni di una scimmia in un test visivo che misura l'acquisizione di un'abitudine.

Dove le memorie si incontrano

Anche l'amigdala ha le sue specializzazioni: molti dei suoi contributi alla memoria sono stati suggeriti dalla sua interessante neuroanatomia, che è stata messa in evidenza molto prima che fosse stabilito il suo ruolo nei processi di memorizzazione. L'amigdala - o, più precisamente, il complesso amigdaloideo, che consiste di numerosi nuclei - è una specie di crocevia nel cervello. Molti ricercatori, fra i quali Blair H. Turner dell'Howard University College of Medicine e uno di noi (Mishkin), hanno dimostrato che l'amigdala presenta connessioni estese e dirette con tutti i sistemi sensoriali corticali. Essa comunica anche con il talamo attraverso una via che è un tratto del sistema mnestico. Infine, le stesse parti dell'amigdala su cui convergono le fibre sensoriali afferenti inviano più profondamente fibre all'ipotalamo, da cui si ritiene abbiano origine le risposte emotive.

La notevole convergenza di connessioni con le aree sensoriali nella amigdala portò Murray e uno di noi (Mishkin) a ipotizzare un suo ruolo nell'associazione di ricordi formati in modalità sensoriali differenti. Fino a quel punto avevamo sempre studiato l'apprendimento in una sua manifestazione, il riconoscimento, che è una risposta all'esperienza visiva, tattile o spaziale. Molto spesso, tuttavia, ciò che risveglia un ricordo è un'esperienza sensoriale di tipo diverso. Udire una voce familiare al telefono ci fa ricordare l'aspetto dell'interlocutore; la vista di una prugna ci riporta alla mente il suo sapore. Per questo meccanismo di richiamo incrociato sembrerebbe necessario un interscambio fra le aree corticali dove i ricordi collegati a ciascuna modalità sensoriale vengono immagazzinati. Questo interscambio potrebbe essere mediato dall'amigdala?

Per mettere alla prova questa ipotesi, abbiamo riunito le versioni visiva e tattile del test di riconoscimento mnestico. Abbiamo fatto eseguire alle scimmie un test di discriminazione ritardata non familiare, che richiedeva di scegliere tra 40 oggetti, ciascuno dei quali aveva caratteristiche visive e tattili diverse. Gli animali eseguivano gli esercizi sia alla luce, sia al buio; in quest'ultimo caso dovevano affidarsi al tatto per distinguere l'oggetto campione da uno nuovo. Dopo la rimozione bilaterale dell'amigdala, le scimmie eseguivano entrambi gli esercizi quasi con la stessa abilità dimostrata prima dell'operazione. La loro memoria di riconoscimento visivo e tattile era rimasta in gran parte intatta, in accordo con le nostre osservazioni precedenti che l'ippocampo e l'amigdala possono sostituirsi l'uno all'altra nella mediazione della memoria di riconoscimento.

Dopo che le scimmie avevano acquisito completa familiarità con le qualità visive e tattili dei 40 oggetti, abbiamo modificato la natura del test. A ogni prova la scimmia doveva prima esaminare l'oggetto nell'oscurità, mediante il tatto, confrontare l'oggetto stesso con un altro alla luce e infine scegliere fra i due affidandosi soltanto alla vista. Per riconoscere l'oggetto che aveva toccato pochi secondi prima, l'animale doveva associare sia ricordi visivi, sia tattili. Un gruppo di controllo costituito da animali cui era stato rimosso l'ippocampo eseguì bene l'esercizio, raggiungendo una precisione del 90 per cento. Animali privati dell'amigdala mostrarono prestazioni di poco superiori al livello di scelta casuale.

La dimostrazione che l'amigdala medii l'associazione di differenti tipi di memoria può chiarire un episodio rimasto misterioso nella storia della neuropsicologia. Circa 50 anni fa, Heinrich Klüver e P. C. Bucy furono colpiti dallo strano comportamento di scimmie i cui lobi temporali erano stati rimossi. Gli animali esaminavano spesso in modo ripetuto e indiscriminato un oggetto non commestibile, toccandolo, assaggiandolo e an-

nusandolo, come se lo trovassero ogni volta estraneo. Lo stesso strano risultato fu osservato, successivamente, in seguito alla rimozione della sola amigdala. In base ai nostri risultati, abbiamo proposto che una delle cause di questo comportamento sia l'incapacità delle scimmie di collegare differenti tipi di memoria. Vedendo un oggetto familiare, esse non possono ricordarne l'odore così come, dopo averlo annusato, non sono in grado di associarlo al suo sapore.

Memoria ed emozioni

Klüver e Bucy notarono un'altra importante caratteristica delle scimmie prive dei lobi temporali, caratteristica che è stata in seguito attribuita alla perdita dell' amigdala. Gli animali non avevano più paura degli esseri umani e perdevano anche la loro avversione per sensazioni normalmente spiacevoli come l'essere pizzicati. Era come se si fosse interrotto un collegamento fra gli stimoli familiari e le loro associazioni emotive. Per le sue connessioni sia con le aree sensoriali corticali, sia con le strutture profonde del cervello che scatenano la risposta emotiva, l'amigdala è una struttura molto adatta a mediare un collegamento di questo tipo. La possibilità che le esperienze sensoriali acquisiscano il loro peso emotivo attraverso l'amigdala è avvalorata dalle osservazioni fatte precedentemente nel corso della nostra indagine sulla memoria: le scimmie private dell'amigdala apprendono lentamente ad associare un oggetto a un premio. Questi animali presentano difficoltà a ricordare le associazioni positive riguardanti uno stimolo familiare.

È possibile che l'amigdala non si limiti a far sì che gli eventi sensoriali sviluppino associazioni emotive, ma consenta anche che le emozioni diano luogo a percezioni e all'immagazzinamento dei ricordi. Come fa il cervello a selezionare gli stimoli significativi in mezzo alla moltitudine di impressioni fornite dai sensi? Se le emozioni possono influenzare l'elaborazione sensoriale che ha luogo a livello corticale, esse devono anche fornire il filtro indispensabile per limitare l'attenzione - e quindi l'apprendimento - a stimoli dotati di significato emozionale. L'amigdala, con la sua capacità di mediare fra i sensi e le emozioni, è una struttura che potrebbe essere alla base di questa «attenzione selettiva».

Esiste una rete nervosa che potrebbe fornire all'amigdala questa funzione di controllo. Numerosi gruppi di ricercatori hanno stabilito che i sistemi sensoriali corticali non solo inviano fibre all'amigdala, ma ne ricevono anche proiezioni, proiezioni che, almeno nel sistema visivo, hanno una densità maggiore nelle stazioni di elaborazione più elevate. In collaborazione con Candace B. Pert e altri, nel suo laboratorio al National Institute of Mental Health, abbiamo trovato un indizio riguardante la natura di alcune di queste proiezioni. L'amigdala è ricca di neuroni che producono alcuni neurotrasmettitori simili all'oppio, gli oppiacei endogeni, che si pensa regolino la trasmissione dei segnali in altre parti del sistema nervoso. Abbiamo trovato che le vie di elaborazione sensoriale a livello corticale mostrano un gradiente di recettori per questi oppiacei: le molecole disposte sulla superficie cellulare alle quali si legano gli oppiacei agendo su un neurone. I recettori sono più abbondanti in corrispondenza delle stazioni finali, dove le impressioni sensoriali si configurano nella loro completezza.

Inoltre, le osservazioni suggeriscono la possibilità che fibre contenenti oppiacei si muovano dall'amigdala verso i sistemi sensoriali, dove possono svolgere una funzione di controllo liberando oppiacei in risposta a stati emotivi generati nell'ipotalamo. In tal modo l'amigdala può fa sì che le emozioni influenzino ciò che viene percepito e appreso. L'effetto reciproco dell'amigdala sulla corteccia può spiegare perché,sia nelle scimmie sia nell'uomo, eventi con una carica emotiva possono restare profondamente impressi.

Memoria e abitudini

Avendo ricostruito due ampi circuiti, uno che trae origine dall'amigdala e l'altro dall'ippocampo, responsabili di molte forme di apprendimento cognitivo - la capacità di riconoscere un oggetto familiare, richiamare le sue qualità sensoriali non percepite, ricordare la sua localizzazione precedente e legare a esso un significato emotivo - ci restava da spiegare perché soggetti che soffrono di una perdita di memoria così completa da essere incapaci di riconoscere una persona vista solo pochi minuti prima sono ancora capaci di apprendere. Alcuni anni fa la Milner riferì che H. M. aveva acquisito la capacità di disegnare allo specchio a una velocità quasi pari a un individuo normale, anche se dopo esservi riuscito non ricordava neppure la prodezza eseguita.

Anche le scimmie le cui strutture limbiche sono state distrutte possono imparare. Le prestazioni di questi animali sono molto scarse allorché vengono sottoposti al test di discriminazione ritardata non familiare, in cui devono riconoscere un oggetto visto soltanto una volta. Tuttavia, nel laboratorio di uno di noi (Mishkin), Barbara L. Malamut ha osservato che, se viene mostrata alle stesse scimmie, una sola volta al giorno, una lunga serie di differenti coppie di oggetti, col tempo esse imparano a scegliere l'oggetto al quale è legato il premio. Inoltre, la loro abilità aumenta a un tasso quasi simile a quello di un animale normale.

TALAMO, TATTO, UDITO, GUSTO, VISIONE, OLFATTO, IPOTALAMO, AMIGDALA

Le molteplici connessioni dell'amigdala spiegano i diversi ruoli che essa svolge nel processo di memorizzazione. Vi sono fibre che raggiungono l'amigdala dalle stazioni terminali dei sistemi corticali sensoriali (*frecce in rosso*); in questo modo le impressioni sensoriali attivano uno dei circuiti del sistema della memoria, quello che dipende dalle connessioni fra amigdala e talamo (*in verde*). Collegamenti fra l'amigdala e l'ipotalamo, dove hanno probabilmente origine le risposte emozionali, permettono a un'esperienza di acquisire un connotato emotivo. Questi collegamenti fanno anche sì che le emozioni influenzino l'apprendimento, attraverso un'attivazione delle connessioni reciproche che vanno dall'amigdala alle vie sensoriali (*in blu, solo per quanto riguarda il sistema visivo*). L'esistenza di connessioni che ritornano alle aree sensoriali dall'amigdala può spiegare perché un singolo stimolo possa suscitare differenti ricordi, come accade quando il profumo di un cibo familiare richiama il suo aspetto, la sua consistenza e il suo sapore.

Nelle stazioni terminali del sistema visivo della scimmia è stato chiaramente individuato un meccanismo molecolare che svolge un ruolo nell'apprendimento. Il grafico, tratto da uno studio condotto da Aryeh Routtenberg della Northwestern University, mostra le quantità relative di fosforo radioattivo incorporato in una proteina nota come *F*1, nel tessuto prelevato da vari punti della via nervosa. Routtenberg ha proposto che la fosforilazione di *F*1 da parte dell'enzima proteinchinasi *C* sia alla base dei cambiamenti che si possono rilevare a livello delle sinapsi di alcuni neuroni, in seguito a stimolazione ripetuta. È possibile che questi cambiamenti riflettano l'immagazzinamento di dati nel cervello. In questo modo, le stazioni visive terminali potrebbero essere adatte, dal punto di vista biochimico, all'immagazzinamento di ricordi.

A un osservatore umano il secondo esercizio sembra più difficile del primo. Come possono conciliarsi questi risultati apparentemente contraddittori?

Come molti altri ricercatori che si occupano dei meccanismi della memorizzazione, uno di noi (Mishkin) ha pensato all'esistenza di un secondo sistema di apprendimento, indipendente dai circuiti limbici. In questo sistema l'elemento critico è la ripetizione stimolo-risposta, proprio quello che non si ritrova nel test di discriminazione ritardata non familiare. In accordo con ciò che si osserva nei pazienti affetti da amnesia, con Herbert L. Petri della Towson State University abbiamo proposto che il secondo sistema dia luogo a un tipo differente di apprendimento,diverso dall'immagazzinamento dei ricordi attraverso il sistema limbico.

Noi chiamiamo questo tipo di apprendimento «abitudine». È un apprendimento non cognitivo: non è fondato sulla conoscenza o anche sul ricordo (nel senso di entità mentali indipendenti), ma su collegamenti automatici fra stimolo e risposta. Nel test di discriminazione fra due oggetti familiari, la scimmia confronta la stessa coppia di stimoli giorno dopo giorno; infine essa sviluppa l'abitudine di prendere l'oggetto la cui scelta è sempre rinforzata dal premio. Il compito di discriminazione ritardata non familiare, invece, non può essere soddisfatto dal formarsi di un'abitudine. Lo stimolo da ricordare viene mostrato una sola volta e la scimmia deve rispondere quindi non allo stesso stimolo - quello prima legato a un premio - ma a uno stimolo nuovo. L'animale deve sapere, in un senso cognitivo, quale degli oggetti è quello originario, in modo da evitarlo.

Le abitudini, come noi le definiamo, sono reminiscenze dei legami automatici stimolo-risposta che gli psicologi comportamentisti ritenevano fossero alla base di tutto l'apprendimento. Il punto di vista comportamentista esclude termini come «mente», «conoscenza» e anche «memoria», come vengono abitualmente intesi. È in contrapposizione con la psicologia cognitiva, che si basa invece su questi concetti per spiegare gran parte del comportamento. La possibilità che l'apprendimento venga elaborato attraverso due sistemi completamente differenti, uno dei quali dia luogo ad abitudini non cognitive, e l'altro sia la base della memoria cognitiva, dà l'opportunità di riconciliare la scuola comportamentista con quella cognitiva. Se esistono meccanismi nervosi per entrambe le forme di apprendimento, il comportamento potrebbe risultare un insieme di risposte automatiche a stimoli e azioni guidate dalla conoscenza e dalle aspettative.

Una struttura adatta per l'instaurarsi di abitudini è il corpo striato, un complesso di strutture localizzato nel cervello anteriore. Il corpo striato riceve proiezioni provenienti da molte aree della corteccia, compresi i sistemi sensoriali, e invia fibre alle parti del cervello che controllano i movimenti. Dal punto di vista neuroanatomico esso è quindi adatto per mettere a disposizione legami relativamente diretti fra stimolo e azione, che sono impliciti nella nozione di abitudine. Altri ricercatori hanno trovato che un danno al corpo striato diminuisce la capacità delle scimmie di contrarre abitudini del tipo di quelle che possono essere valutate con il test di discriminazione fra due oggetti familiari.

Paul D. MacLean del National Institute of Mental Health ha sottolineato che il corpo striato è una parte evolutivamente antica del cervello, molto più vecchia della corteccia e del sistema limbico. In effetti, l'acquisizione di abitudini dovrebbe essere mediata da strutture primitive: anche gli animali più semplici possono apprendere risposte automatiche a stimoli. Dal punto di vista dello sviluppo, l'abitudine sembra anch'essa primitiva. La Bachevalier ha scoperto recentemente che scimmie giovani si comportano come quelle adulte nel nostro test di formazione di abitudini e che mostrano prestazioni scarse nel test di memoria. Rispetto ai criteri utilizzati per l'adulto esse soffrono di amnesie. Stiamo ora vagliando la possibilità che il substrato nervoso dell'abitudine sia pienamente sviluppato nelle scimmie giovani, mentre il sistema della memoria presenti una maturazione più lenta. La stessa differenza evolutiva, se fosse presente negli esseri umani, potrebbe spiegare perché poche persone conservano ricordi precisi della loro infanzia.

Per quel che riguarda l'interazione fra memoria e abitudine nel cervello maturo, stiamo appena incominciando a porci delle domande. Sembra probabile che la maggior parte dell'apprendimento si basi su entrambi i sistemi, ma è facile immaginare che la memoria cognitiva e l'abitudine non cognitiva possano entrare spesso in conflitto. Come può il cervello decidere fra la formazione di un'abitudine e l'apprendimento cognitivo? Gli elementi del sistema della memoria comunicano con il corpo striato influenzando quindi la formazione di abitudini? Per ora siamo riusciti solamente a tracciare una mappa di un territorio cerebrale che è ancora tutto da esplorare.

Crollo di vulcani

Lo studio dei depositi formati da enormi valanghe di detriti rocciosi rivela le modalità secondo le quali si verifica il crollo catastrofico dell'edificio vulcanico: un evento non infrequente nel ciclo vitale di queste montagne

di Peter Francis e Stephen Self

I vulcani sono montagne particolari. La maggior parte delle montagne deve la propria forma soprattutto all'erosione per opera delle piogge e dei ghiacci, che scolpiscono senza posa i grandi blocchi di crosta sollevati dall'attività tettonica. Gli edifici vulcanici, invece, si autocostruiscono con la lava e le ceneri provenienti dall'interno della Terra. Poiché i vulcani possono aumentare la propria altezza per aggiunta di nuovi materiali, ci si potrebbe chiedere per quale ragione essi non siano invariabilmente le montagne più alte sulla Terra (come avviene in effetti su Marte e probabilmente anche su Venere). Quali sono i fattori che limitano l'altezza di un vulcano sulla Terra?

Ovviamente anche i vulcani sono soggetti all'erosione e ad altri processi, fra cui la flessione della litosfera sotto il peso delle masse rocciose, che tendono a livellare la topografia terrestre. Negli ultimi anni, però, è diventata chiara l'esistenza di un altro processo, assai più drammatico, che limita l'altezza di molti edifici vulcanici: il loro collasso, che si manifesta sotto forma di crollo catastrofico. Sebbene tale fenomeno sia stato osservato varie volte nel secolo scorso, la terrificante eruzione del Mount St. Helens del 1980 richiamò l'attenzione sull'eventualità che l'intero fianco di un vulcano possa di colpo crollare, innescando non solo un'eruzione esplosiva, ma anche una vera e propria valanga devastante di detriti. Nel tentativo di comprendere il fenomeno, vari ricercatori hanno rivolto la loro attenzione ai depositi delle valanghe di detriti accumulati sia presso il Mount St. Helens, sia presso altri vulcani. Finora sono stati identificati più di 100 di tali depositi, la cui età varia da alcuni anni a decine di migliaia di anni e le cui dimensioni vanno da meno di uno a più di 20 miliardi di metri cubi.

Nella nostra ricerca, abbiamo esaminato immagini Landsat per identificare depositi di detriti in precedenza sconosciuti attorno a vulcani nelle Ande centrali. Inoltre abbiamo studiato dettagliatamente sul campo il deposito più imponente, che si trova sulle pendici di un vulcano cileno, il Socompa. Grazie a questo lavoro e a quello di molti altri ricercatori, in particolare di Tadahide Ui dell'Università di Kobe, sta diventando chiaro che il crollo catastrofico di una parte instabile del proprio edificio può essere considerato un evento normale nel ciclo vitale di un vulcano. Questa conclusione è valida particolarmente per i grandi stratovulcani, molte centinaia dei quali si trovano nella «cintura di fuoco» attorno al Pacifico.

Il Socompa, un vulcano del Cile settentrionale che circa 7500 anni fa crollò dando origine a una grande valanga di detriti, è stato ripreso dal *thematic mapper*, un'apparecchiatura trasportata a bordo del Landsat. L'immagine in falsi colori è una composizione di dati raccolti a varie lunghezze d'onda. Nel deposito del materiale franato, che si estende su circa 600 chilometri quadrati, si distinguono, in base alle caratteristiche spettrali, varie correnti di detriti. Le zone grige a nord e a est sono soprattutto ghiaie e ignimbriti del basamento vulcanico; nelle zone rossastre, questi depositi sono stati ricoperti da lave provenienti dal crollo del versante del vulcano. Le zone blu sono nevi della vetta del Socompa, che è alto 6051 metri.

Una delle ragioni per cui gli stratovulcani sono particolarmente soggetti al fenomeno ha attinenza con la loro composizione. Essi sono generalmente associati a una zona di subduzione: la regione in cui due zolle litosferiche convergono e una delle due si immerge sotto l'altra. (La cosiddetta «cintura di fuoco» è in effetti un anello di zone di subduzione.) Un vulcano sopra una zona di subduzione tende a emettere lave andesitiche o dacitiche, che sono relativamente ricche di silice e perciò viscose. Anziché scorrere, le lave si accumulano attorno alla bocca del vulcano, formando un cono ripido che può diventare meccanicamente instabile.

Si possono distinguere tre tipi di collasso. Il primo è rappresentato dall'evento del Mount St. Helens, dove il processo innescò una violenta esplosione. L'intrusione di una massa di nuovo magma determinò la formazione di una grande protuberanza che rese instabile il fianco settentrionale della montagna; il magma interagì con le acque sotterranee generando vapore surriscaldato, che rimase intrappolato ad alta pressione all'interno del vulcano. Quando un terremoto provocò il crollo parziale del cono, il vapore sfuggì con violenza esplosiva, proiettando tutt'attorno ceneri e detriti.

Il secondo tipo di collasso è rappresentato dal vulcano Bandai-san in Giappone. Benché il suo crollo nel 1888 sia stato innescato da un terremoto e accompagnato da esplosioni, non c'è alcuna prova che nuovo magma si fosse intruso nel vulcano. Le esplosioni furono probabilmente la conseguenza dell'interazione di acque sotterranee con una grande massa

Il collasso del Socompa avvenne in quattro fasi distinte, che nel loro insieme non durarono, probabilmente, più di dieci minuti. Il vulcano era stato destabilizzato dall'intrusione di nuovo magma, che potrebbe aver formato duomi di lava sul suo versante occidentale. La catastrofe fu probabilmente innescata da un terremoto lungo la scarpata di faglia sotto il fianco occidentale (*1*). La prima superficie di distacco correva alla base della montagna (*sezione trasversale*), cosicché la prima ondata di materiale di valanga (*in grigio*) era costituita soprattutto da ghiaie e ignimbriti del basamento del vulcano. I detriti si mossero verso nord-ovest, trasversalmente alla pendenza locale, a una velocità di forse 300 chilometri all'ora. Ancor prima di raggiungere il suo punto più avanzato, a circa 35 chilometri dalla vetta del Socompa, questa prima corrente di detriti fu raggiunta e superata da una seconda ondata (*in arancione*), messa in movimento dal crollo dell'edificio vulcanico stesso (*2*) che conteneva un'alta proporzione di nuovo magma. Le due correnti di detriti si fermarono pochi istanti, poi gran parte del materiale cominciò a scivolare verso nord-est, seguendo la lieve pendenza locale (di tre gradi) (*3*). Nello stesso tempo, un altro crollo presso la vetta del vulcano diede origine a una terza massa di materiale detritico (*in giallo*) che scese a sud-ovest delle prime due. Parti delle correnti di detriti continuarono a scendere a valle, mescolandosi con i materiali della valanga primaria e lasciando un'evidente scarpata marginale (*4*). Infine, grandi blocchi di roccia si staccarono dalle pareti dell'anfiteatro formato dal collasso e scivolarono intatti per chilometri fino al suo sbocco.

di roccia molto calda, ma non più fluida.

Infine, alcuni eventi di questo genere si sono verificati senza essere accompagnati da esplosioni e attività vulcanica; il vulcano franò semplicemente con una valanga «fredda», simile a quelle che si osservano in montagne non vulcaniche. L'esempio più recente si verificò al vulcano Ontake, in Giappone, nel 1984.

Il collasso del Mount St. Helens è l'unico che sia mai stato osservato da vicino. Come si può riconoscere un evento avvenuto in epoca preistorica? Nei pochi casi in cui la valanga di detriti pose drammaticamente fine alla vita del vulcano il compito è semplice: il vasto anfiteatro a forma di ferro di cavallo lasciato dal movimento delle masse rocciose rimane visibile per migliaia di anni. Nel caso di un vulcano attivo, invece, la cicatrice può essere cancellata da nuove effusioni di lava. Al Mount St. Helens la valanga strappò dal fianco dell'edificio vulcanico circa 2,8 miliardi di metri cubi di roccia, creando un gigantesco anfiteatro della larghezza di due chilometri e della profondità di 600 metri, ma oggi un duomo di lava in accrescimento occupa già circa un terzo della sua base. Gli anfiteatri del Tata Sabaya in Bolivia e del Parinacota in Cile, che subirono entrambi crolli catastrofici migliaia di anni fa, sono stati da allora cancellati completamente da flussi e duomi di lava.

In questi casi è necessario individuare i depositi lasciati dalla valanga di detriti vulcanici. La loro identificazione non è sempre facile: essi possono essere occultati da suolo più recente e dalla vegetazione, oltre a poter essere scambiati per il prodotto di normali processi di erosione e deposizione. Nonostante ciò, la struttura caratteristica di un deposito di valanga è spesso ben riconoscibile. L'indizio più evidente è la presenza di morfologie tormentate: migliaia di piccoli poggi e depressioni che possono coprire decine o anche centinaia di chilometri quadrati alla base di un vulcano. Una morfologia di questo genere è stata la caratteristica che più ci ha guidato nell'identificazione di questi depositi, quando abbiamo esaminato fotografie prese da satellite alla ricerca di crolli vulcanici antichi nelle Ande centrali.

L'esempio più spettacolare fra quelli da noi scoperti è il Socompa, un vulcano che si trova nel Cile settentrionale, al confine con l'Argentina, e che subì un enorme crollo circa 7500 anni fa. Mentre l'anfiteatro del Mount St. Helens occupa circa 30 gradi della circonferenza del vulcano, al Socompa la valanga portò via un cuneo di 70 gradi e il volume del materiale che spazzò il fianco della montagna fu dieci volte maggiore di quello del Mount St. Helens. Il deposito della valanga, con una superficie di circa 600 chilometri quadrati, è uno dei più grandi nel mondo ed è conservato magnificamente. Esso fa del Socompa un caso esemplare di collasso vulcanico.

Attualmente il Socompa è alto 6051 metri, ma prima della catastrofe aveva probabilmente un'altezza di circa 6300 metri. La montagna è molto più alta sul lato occidentale - dove si innalza di quasi 3000 metri al di sopra del Deserto di Atacama - che non sul versante orientale, dove il dislivello fra la base e la vetta è di circa 2000 metri. La differenza di altitudine fra il livello della base a oriente e a occidente è un indizio importante per determinare la causa del collasso. Il terreno a est del Socompa si trova a una quota più alta perché è stato sollevato lungo una serie di importanti faglie di distensione con direzione nord-sud che interessano il fianco occidentale della montagna; è proprio qui, in effetti, che il crollo ebbe inizio. Ovviamente, dove ci sono faglie si verificano di solito terremoti e sembra molto probabile che un forte terremoto lungo una delle faglie abbia innescato il crollo del Socompa, esattamente come avvenne nel caso del Mount St. Helens. Tracce delle faglie sono tuttora visibili; esse tagliano depositi di ghiaia recenti, il che induce a pensare che possano essere ancora attive a 7500 anni di distanza dal cataclisma.

I blocchi «a crosta di pane» nel deposito dei detriti di valanga del Socompa dimostrano che nel vulcano, prima del suo crollo, era presente nuovo magma. Lo strato esterno vetroso e l'interno schiumoso del blocco in primo piano indicano che esso si raffreddò rapidamente dallo stato fuso durante la discesa della valanga. Le ondulazioni in secondo piano sono i blocchi di scorrimento, alti diverse centinaia di metri e lunghi fino a due chilometri, staccatisi dal Socompa nella fase finale del collasso. La vetta del Socompa si trova a 16 chilometri di distanza.

In realtà, il Socompa dev'essere stato scosso da decine e decine di terremoti per migliaia di anni prima del suo collasso. Perché allora non crollò prima? Anche qui la risposta sembra essere la stessa che nel caso del Mount St. Helens. Le osservazioni effettuate sul luogo dimostrano che nuovo magma doveva essere risalito all'interno del Socompa e potrebbe anche essere stato estruso sotto forma di lava nella parte alta del cono. Il deposito di valanga comprende numerosi blocchi a «crosta di pane», così denominati perché il loro strato superficiale vetroso e la loro struttura interna schiumosa li fanno assomigliare a pane ben cotto, con una crosta consistente. Alcuni di questi blocchi hanno dimensioni che possono raggiungere i 20 metri. Essi avrebbero potuto essere originati solo da una nuova massa di magma che, intrudendosi nel vulcano, avrebbe contribuito a renderlo instabile.

Secondo la nostra opinione, fu la fatale coincidenza di un terremoto con un periodo di attività magmatica a condurre al collasso del Socompa, mentre è incerto se il crollo abbia o meno innescato a sua volta un'esplosione. Nel caso del Mount St. Helens, il caratteristico deposito di ceneri e detriti lasciato dalla grande esplosione è estremamente sottile (il suo spessore varia da alcuni centimetri a un metro circa) e nell'arco di pochi decenni sarà in gran parte asportato dall'erosione. Non sorprende, quindi, che non ci siano prove sicure di un'esplosione al Socompa. In generale, è comunque molto probabile che una grande valanga vulcanica inneschi un'eruzione esplosiva ogni qual volta nel vulcano vi sia magma. Lo sventramento del vulcano porta a un'improvvisa diminuzione della pressione nel suo interno caldo, una condizione propizia per un'esplosione.

La valanga del Socompa portò al distacco di una grossa porzione della montagna. Mentre la maggior parte dei detriti del Mount St. Helens proviene dall'edificio vulcanico stesso, una frazione considerevole (probabilmente più di metà) dell'enorme massa di materiale franato del Socompa derivò non dal vulcano vero e proprio, bensì dal basamento sottostante. Questo, che si estende sotto la quota di circa 4000 metri, è composto di rocce vulcaniche molto più antiche del cono e di altre rocce che non sono di origine vulcanica. Perciò, nel deposito di valanga del Socompa si trovano non soltanto i soliti blocchi di lava, ma anche molte ignimbriti (rocce formatesi dalle ceneri consolidate di antiche eruzioni) e ghiaie sedimentarie. Noi riteniamo che queste rocce si siano distaccate perché la superficie di scivolamento attraversò la faglia occidentale, incidendo profondamente il basamento. I materiali franati per primi provenivano soprattutto dal basamento, non dalle parti superiori della montagna.

L'energia della valanga fu straordinaria. Quando la prima ondata di materiali franosi, che procedeva verso nord-ovest, raggiunse la base del Socompa, non deviò verso nord-est, seguendo la pendenza locale, ma proseguì il suo cammino, in linea retta, fino a fermarsi contro un ripido pendio a circa 35 chilometri di distanza dalla vetta del vulcano. (Il limite nordoccidentale della valanga è contrassegnato da un fronte alto più di 40 metri.) Il fatto che i materiali franati non abbiano seguito una traiettoria curva - come quella di una palla che rotoli lentamente attraversando un piano inclinato - suggerisce che stessero muovendosi a grande velocità. Non vi è modo di determinarne direttamente la velocità ma, a giudicare da valanghe effettivamente osservate, essa era probabilmente dell'ordine di 300 chilometri orari.

Il crollo del basamento che stava sotto il vulcano determinò il franamento dell'edificio vulcanico stesso, con il risultato che la prima ondata della valanga fu quasi immediatamente raggiunta e superata da una seconda ondata proveniente da una zona più alta della montagna, che si muoveva anch'essa verso nord-ovest. Le due masse di detriti, giunte in basso, si

Il Parinacota, un altro vulcano che si trova nel Cile settentrionale, subì un collasso tra 15 000 e 17 000 anni fa. La fotografia mostra due caratteristiche dei depositi di valanga vulcanica: sulla sinistra, ai piedi del vulcano, si vedono grandi blocchi di scorrimento; davanti a essi vi sono collinette più piccole. Un lago riempie le depressioni fra i rilievi. A giudicare dalle dimensioni del deposito, che si estende per circa 150 chilometri quadrati, il collasso del Parinacota fu meno catastrofico di quello del Socompa. L'anfiteatro è stato cancellato da colate laviche.

fermarono solo per breve tempo. Poi tutto il materiale che si trovava a una distanza superiore ai 25 chilometri circa dalla vetta, eccezion fatta per una fascia ristretta lungo il margine nordoccidentale della valanga, cominciò a scivolare ad angolo retto rispetto al percorso iniziale, seguendo la pendenza locale verso nord-est. Questa corrente secondaria si sovrappose al materiale della prima valanga, perciò il suo margine rivolto verso la montagna presenta un'evidente scarpata, o dirupo, con andamento verso nord-est. Successivamente, una terza ondata di materiale, proveniente dalla zona sommitale del vulcano, ricoprì parzialmente i primi due flussi di detriti.

In definitiva una parte del materiale caduto dal Socompa percorse quasi 40 chilometri prima di fermarsi nel deserto. Come può aver percorso una distanza tanto grande? Si potrebbe pensare che sia stata scagliata lontano da un'enorme esplosione, ammesso che un'esplosione si sia realmente verificata al Socompa. In realtà, persino esplosioni di grande potenza non sono in grado di fornire un contributo significativo all'energia di una valanga vulcanica. La sua energia, analogamente a quella di una valanga comune, dipende soprattutto dalla forza di gravità; in altri termini, più il punto di partenza è posto in alto, maggiore è la distanza percorsa. Il ripido pendio occidentale del Socompa contribuisce perciò a spiegare perché la valanga si sia spinta così avanti nel Deserto di Atacama.

Ma la ripidità del pendio non è una spiegazione sufficiente. La valanga del Socompa percorse una distanza molto maggiore di quella che ci si potrebbe attendere dalla topografia della regione. Gran parte del materiale proveniva dal basamento del vulcano, il che significa che prese l'avvio da un'altitudine di 4000 metri o anche meno. Il margine nordoccidentale del deposito si trova a una trentina di chilometri dalla base della montagna, a un'altitudine di 3200 metri. Dividendo il dislivello verticale per la distanza orizzontale percorsa si ottiene un rapporto di meno di 0,03. (Il calcolo per il movimento franoso secondario, in direzione nord-est, fornisce un risultato sorprendentemente simile.) Il rapporto tra la distanza verticale e quella orizzontale è un buon indicatore della mobilità di una valanga; un valore tipico è circa 0,1. Sembra quindi che la valanga del Socompa sia stata insolitamente mobile.

La spiegazione più plausibile del fenomeno può essere trovata nel fatto che il terremoto che innescò il collasso del vulcano deve anche aver messo in movimento la valanga. Un terremoto di grande energia (di magnitudine sette o otto nella scala Richter), può continuare a scuotere intensamente il terreno per parecchi minuti; inoltre, le scosse possono venire intensificate e prolungate quando le onde sismiche rimbalzano avanti e indietro a causa della configurazione a bacino delle rocce subsuperficiali (come quella presente al Socompa). La valanga del Socompa deve aver impiegato solo pochi minuti per raggiungere il suo punto più lontano. L'apporto di energia sismica deve averla mantenuta in moto, permettendole di superare la perdita di energia dovuta all'attrito. In particolare, le scosse sismiche devono aver rimesso in movimento i materiali della valanga primaria dopo la breve sosta, dando così inizio alla corrente secondaria di detriti lungo la pendenza locale, meno ripida.

Il quadro che abbiamo presentato - quello di una valanga catastrofica, rimescolata da scosse violente mentre scende lungo il fianco di una montagna a velocità tremenda - può dare l'impressione che il flusso di detriti risultante da un collasso vulcanico sia altamente disorganizzato e turbolento. In realtà non è così. Il flusso è in effetti complesso, soprattutto una volta che abbia raggiunto la velocità massima, ma non è caotico o turbolento. La caratteristica di un flusso turbolento, che si osserva principalmente in fluidi di bassa viscosità come l'aria o l'acqua, è che le singole particelle nella massa che fluisce sono costantemente in moto l'una rispetto all'altra. Al contrario, il flusso in una valanga vulcanica tende a essere laminare, ossia a procedere per strati piani che conservano le proprie posizioni relative. Soltanto lo strato più in basso, quello a contatto con il terreno, viene sottoposto a notevoli sforzi di taglio.

L'anfiteatro del Mount St. Helens ha una larghezza di due chilometri e una profondità di 600 metri. La fotografia, scattata nel settembre 1984 da Lyn Topinka dell'US Geological Survey, mostra il duomo di lava formatosi dopo le eruzioni, che occupa un terzo del fondo dell'anfiteatro.

Il vulcano Bandai-san, in Giappone, crollò il 15 luglio 1888. Questa stampa d'epoca, che si basa su fotografie e disegni fatti tre settimane dopo il disastro, illustra con precisione l'anfiteatro pieno di detriti (con diverse bocche vulcaniche fumanti) e la morfologia tormentata dei depositi di valanga. A differenza del Mount St. Helens, sembra che il Bandai-san non fosse stato destabilizzato dalla risalita di nuovo magma.

In effetti, il moto della valanga nella prima parte del suo percorso, sul fianco del vulcano, può essere descritto più accuratamente come uno scivolamento piuttosto che come un flusso. Blocchi enormi di roccia si staccano dalla montagna e accelerano mentre precipitano a valle. Nonostante una certa rotazione all'indietro, i blocchi rimangono intatti. Durante questa fase, sarebbe forse possibile scendere a valle senza eccessivi pericoli stando sulla superficie di uno di questi blocchi. Nel caso del Socompa, alcuni blocchi - quelli che si sono staccati dalla zona della vetta, in prossimità dell'estremo superiore della valanga - si ritrovano intatti allo sbocco dell'anfiteatro, a vari chilometri di distanza dal loro punto di partenza. Il più grande di essi è lungo all'incirca due chilometri e alto mezzo chilometro.

Senza dubbio, grandi blocchi di roccia si staccano dal vulcano per tutta la durata di una valanga, ma i primi non si arrestano allo sbocco dell'anfiteatro e non si conservano intatti. Quando un blocco arriva ai piedi del vulcano, ha una velocità compresa fra 100 e 200 chilometri all'ora. Gli sforzi di taglio, a tali velocità, sono sufficienti a frantumare il blocco in frammenti minori, la maggior parte dei quali di diametro inferiore a un metro. Il movimento di queste rocce nel loro allontanarsi dal vulcano è più un flusso che uno scorrimento, ma non è turbolento; nel deposito di detriti, si mantengono alcune correlazioni stratigrafiche presenti nel vulcano.

Quest'osservazione è stata fatta su vari depositi di valanga e può essere verificata su scala eccezionalmente grande al Socompa: nel vulcano, lave dacitiche di colore chiaro poggiavano sopra lave andesitiche scure e nel deposito i frammenti di dacite si possono ritrovare ancora sopra quelli di andesite. Al di sotto di esse, particolarmente al margine più esterno del deposito, si possono vedere ghiaie e ignimbriti di colore chiaro provenienti dal basamento subvulcanico. Tutti e tre gli strati sono stati sottoposti a un massiccio assottigliamento a causa dell'attrito. Mentre in origine avevano uno spessore di decine o centinaia di metri, nei depositi di valanga il loro spessore si è ridotto a non più di qualche metro.

Nulla può sopravvivere lungo il percorso di una valanga che procede ad alta velocità e che, a causa della sua enorme quantità di moto, è in grado di risalire un pendio per centinaia di metri. Oltretutto, la distruzione può estendersi molto al di là delle zone investite dai detriti perché, sconvolgendo il sistema idrografico regionale e provocando lo straripamento delle acque di fiumi e laghi, una valanga può generare grandi colate di fango. Peggio ancora, se il vulcano sorge in prossimità di un oceano o di un grande lago, l'improvviso spostamento d'acqua a opera dei materiali caduti può dare origine a uno tsunami, una gigantesca ondata dagli effetti distruttivi. Il crollo del cono Rakata su Krakatoa, nel 1883, è un esempio tristemente famoso. Quando i detriti del cono precipitarono in mare al culmine dell'eruzione magmatica, provocarono la «grande ondata» che devastò per centinaia di chilometri le coste nello Stretto della Sonda (fra Giava e Sumatra), spazzando via intere città e provocando la morte di circa 30 000 persone.

La maggior parte dei collassi vulcanici è probabilmente preceduta da attività magmatica e accompagnata da un'eruzione esplosiva. In tali casi la forma del vulcano può consentire di prevedere un possibile cataclisma. Prima dell'eruzione del Mount St. Helens, per esempio, i vulcanologi che stavano controllando la crescita della protuberanza sul suo fianco settentrionale si resero perfettamente conto che la montagna poteva crollare (anche se adesso è facile dire che sottovalutarono l'entità del disastro). Quando un vulcano fornisce segni evidenti di un prossimo collasso, è possibile preparare carte con l'indicazione delle zone che rischiano di subire maggiori danni dall'esplosione o dalla valanga. Per ragioni comprensibili, i ricercatori che studiano questi rischi tendono a concentrare la loro attenzione su vulcani con segnalazioni di attività nel recente passato.

È nondimeno del tutto possibile che, in conseguenza di un grande terremoto, un vulcano inattivo subisca un collasso senza aver dato alcun segno premonitore: una «valanga fredda» di questo genere può essere altrettanto distruttiva di una grande eruzione. Nel 1792, per esempio, il fianco del vulcano Unzen, nel Giappone meridionale, crollò, precipitando nella Baia di Ariake, e il successivo tsunami uccise oltre 14 500 persone. Indubbiamente, questi eventi costituiscono un'ulteriore minaccia per gli abitanti delle coste dell'oceano Pacifico, dove vulcani e terremoti sono numerosi.

Gli adattamenti all'immersione della foca di Weddell

Grazie alla comprimibilità dei suoi polmoni e alla milza che funge da «autorespiratore», questa foca è in grado di nuotare a grandi profondità e di trattenere il respiro più a lungo della maggior parte dei mammiferi

di Warren M. Zapol

Un subacqueo in grado di immergersi fino a una profondità di 20 metri e di rimanere in apnea per 3 minuti è considerato esperto. Eppure una prestazione del genere non è neppure lontanamente paragonabile a quella di un altro mammifero, che è capace di raggiungere in immersione i 500 metri di profondità e di restare sott'acqua per più di 70 minuti. Questo virtuoso dell'immersione è la foca di Weddell (*Leptonychotes weddelli*), un membro della famiglia *Phocidae*, le foche meglio adattate alla vita acquatica, come dimostra il fatto che sono prive di padiglioni auricolari.

Quest'animale, molto diffuso lungo i litorali e i margini della banchisa dell'Antartide, si tuffa nelle gelide acque dell'oceano non certo nel tentativo di stabilire un record, ma per procurarsi il cibo. A un quarto di miglio dalla terraferma e a una profondità compresa tra i 250 e i 600 metri vive infatti la sua principale risorsa alimentare: il grande merluzzo antartico *Dissostichus mawsoni*.

Le foche di Weddell sopportano con facilità temperature dell'acqua che possono raggiungere −1,9 gradi Celsius grazie alle loro notevoli dimensioni (un adulto pesa da 350 a 450 chilogrammi) e a uno spesso rivestimento di grasso sottocutaneo che funge da isolante. L'immersione, che costringe l'animale a far fronte alla mancanza d'aria e alla forte pressione idrostatica, costituisce una sfida fisiologica più complessa. Le indagini sugli adattamenti sviluppati in risposta a questa necessità hanno richiesto decenni di ricerche di laboratorio da parte di un buon numero di studiosi e, più recentemente, parecchi studi «sul campo», o meglio, in mare. Questi ultimi fanno pensare che sia opportuno rivedere alcune opinioni, accettate da tempo, basate su esperienze di laboratorio. In effetti, costringere una foca tenuta in cattività a mettere la testa sott'acqua non provoca necessariamente la stessa risposta che si avrebbe in un'immersione effettuata spontaneamente in mare.

I problemi specifici posti dall'immersione sono considerevoli. Innanzitutto, la foca deve rifornire i propri tessuti di ossigeno e, nello stesso tempo, deve limitare l'accumulo nel sangue di anidride carbonica, un sottoprodotto del processo di ossidazione del glucosio che fornisce energia all'animale. Questo gas si forma nei tessuti e successivamente viene trasportato dal flusso sanguigno ai polmoni, dove viene eliminato. Quando un animale è in immersione, l'anidride carbonica si accumula nel sangue, alterandone il delicato equilibrio del *p*H.

L'animale, inoltre, deve evitare i numerosi disturbi che dipendono dalla forte pressione. Infatti, ogni 10 metri di profondità, la pressione idrostatica che grava su un corpo in immersione aumenta di un'atmosfera, che è la spinta esercitata a livello del mare da una forza pari a un chilogrammo-peso per centimetro quadrato, ovvero la pressione esercitata da una colonna di mercurio alta 760 millimetri. Un possibile effetto dell'aumento di pressione è una maggiore eccitabilità delle cellule nervose, che può condurre a episodi convulsivi. Inoltre, la pressione tende a comprimere gli spazi pneumatici del corpo, come i seni nasali o frontali nel cranio umano. Lo schiacciamento può provocare una sensazione di dolore e, se il corpo non ha possibilità di introdurre in tali spazi aria sufficiente per equilibrare la pressione esterna, i vasi sanguigni possono espandervisi, fino a scoppiare.

L'aumento di pressione porta alla compressione dei gas presenti nell'organismo, un fenomeno che si rivela pericoloso quando interessa l'azoto che si trova negli alveoli polmonari. (I fluidi corporei e gli organi ricchi di liquidi, al contrario, subiscono una compressione minima.) L'azoto è il più abbondante dei gas atmosferici (78 per cento). Di solito esso transita nell'apparato circolatorio senza recar danno, ma, quando l'aria presente nei polmoni viene sottoposta a un'elevata pressione, come si verifica durante un'immersione, l'azoto che risulta in eccesso rispetto alle condizioni atmosferiche si scioglie nel sangue e nei tessuti e può condurre a narcosi, un grave disturbo chiamato anche «ebbrezza da profondità». I sintomi della narcosi sono: avvelenamento del sangue, perdita della coordinazione motoria, alterazione della percezione visiva, sonnolenza e svenimento. Dopo la permanenza in immersione, una risalita troppo rapida può far sì che la pressione parziale dell'azoto nel sangue e nei tessuti risulti maggiore della pressione esercitata dall'ambiente esterno sul corpo. L'azoto disciolto può allora liberarsi dalla soluzione, sviluppando bollicine, che oltre a provocare dolori alle articolazioni e in altre parti del corpo, possono ostruire i vasi sanguigni nel cervello e nel midollo spinale, portando a paralisi e anche alla morte per embolia gassosa.

Gli esperimenti svolti in laboratorio, pur con i loro limiti, hanno chiarito molte strategie che consentono alle foche di assicurarsi un'adeguata riserva di ossigeno e di evitare i disturbi citati. In questo articolo esaminerò tali strategie e quindi segnalerò i casi in cui i risultati degli studi sul campo hanno contraddetto o chiarito quelli ottenuti in laboratorio.

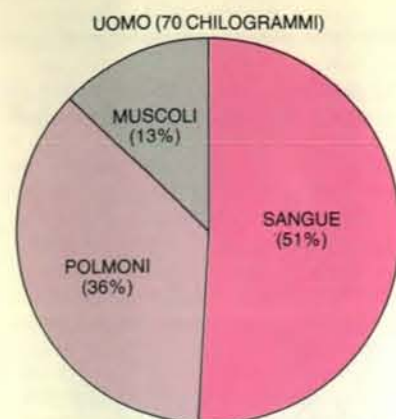
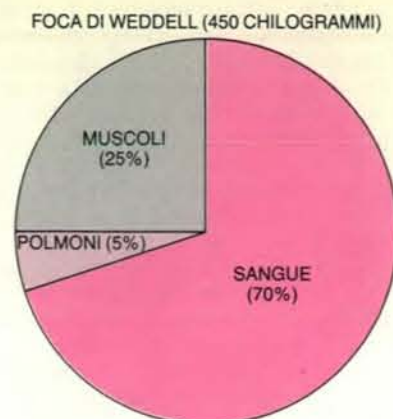
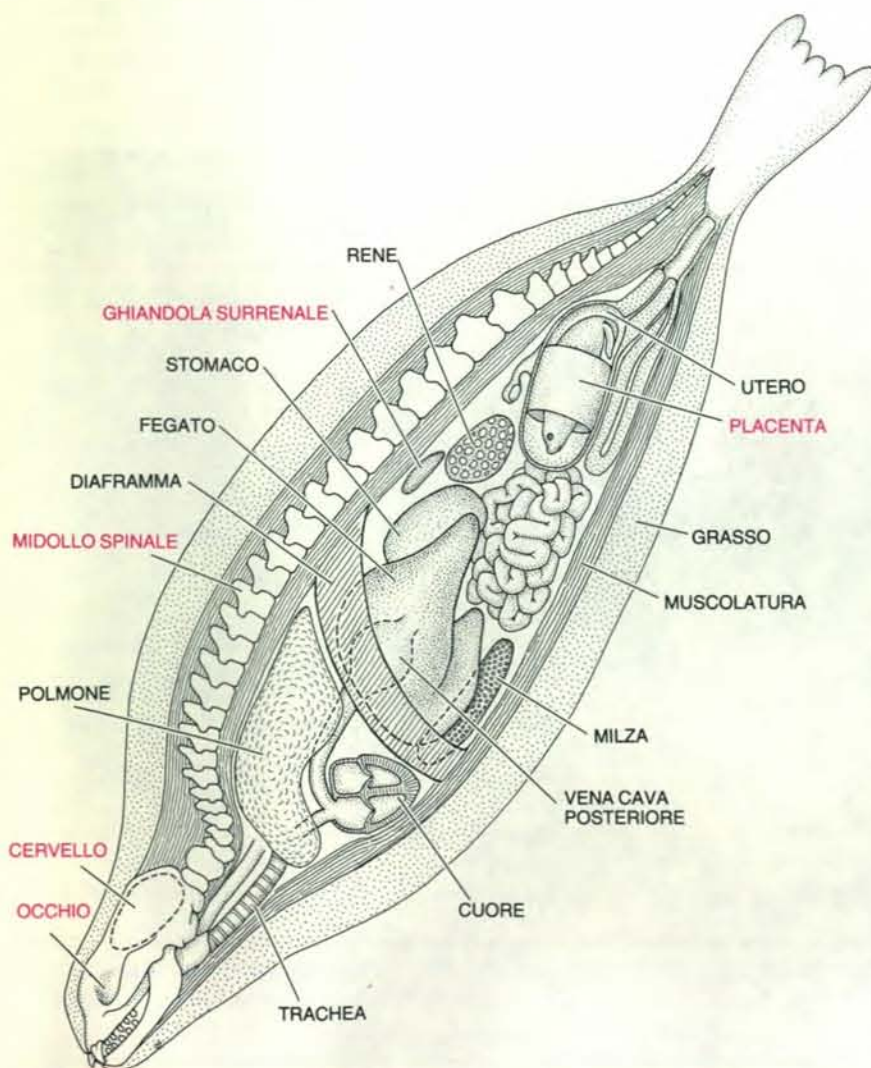
Una delle più importanti e ancora indiscusse scoperte di laboratorio è che la foca immagazzina ossigeno in grande quantità, quasi il doppio, per chilogrammo di peso, rispetto all'uomo. Essa, inoltre, concentra l'ossigeno proprio dove è più necessario durante l'immersione: nel sangue e, in misura minore, nei muscoli. L'uomo dipende strettamente dalla riserva polmonare di ossigeno, in quanto il 36 per cento della sua riserva totale è contenuto nei polmoni e il 51 per cento nel sangue, mentre la foca immagazzina solo il 5 per cento dell'ossigeno nei polmoni e il 70 per cento nel sangue. Analogamente, la foca di Weddell conserva il 25 per cento dell'ossigeno nei muscoli (rispetto al 13 per cento soltanto nell'uomo), legato alla mioglobina, un pigmento trasportatore di ossigeno. (La mioglobina è presente anche nell'uomo, ma in quantità inferiore.)

Il volume di sangue della foca è enorme e questo spiega in parte perché possa contenere una quantità così elevata di ossigeno. Nel 1969, Claude J. M. Lenfant, allora all'Università di Washington, scoprì che, a differenza dell'uomo, il cui sangue rappresenta normalmente il 7 per cento del peso corporeo, nella foca di Weddell il peso del sangue è pari al 14 per cento del peso dell'animale. (Questo confronto, in realtà, sottovaluta la quantità di sangue disponibile per i tessuti che compiono lavoro, in quanto il grasso, che costituisce circa un terzo della massa dell'animale, è poco irrorato di sangue.) Per di più, il sangue della foca contiene grandi quantità di emoglobina, il pigmento trasportatore di ossigeno dei globuli rossi. Quando i miei colleghi e io al Massachusetts General Hospital abbiamo prelevato campioni di sangue da foche tenute in laboratorio, abbiamo visto che i globuli rossi rappresentavano circa il 60 per cento del volume di ogni goccia; nell'uomo essi occupano solo dal 35 al 45 per cento di tale volume.

Sebbene la riserva di ossigeno della foca di Weddell sia enorme, non è tuttavia illimitata. Come altri animali che compiono lunghe immersioni, la foca ha perciò sviluppato alcuni adattamenti per conservarla il più a lungo possibile. Quando un mammifero mette la testa sott'acqua, al cervello arrivano impulsi nervosi che inducono il riflesso da immersione: non appena l'animale cessa di respirare, insorge la bradicardia (cioè il rallentamento del battito cardiaco) e alcune arterie si contraggono, limitando così l'afflusso di sangue agli organi che esse irrorano.

La rapida insorgenza della bradicardia all'inizio di un'immersione è conosciuta negli animali da più di un secolo. Essa si manifesta anche nell'uomo, ma sembra essere più accentuata in quelle specie che si immergono abitualmente, come le foche e le balene. Sott'acqua, un rallentamento del battito cardiaco è utile, perché consente al cuore di lavorare meno e quindi di consumare una minor quantità di ossigeno. La bradicardia riduce anche la gettata cardiaca (il flusso di sangue in uscita dal cuore), il che contribuisce a mantenere la pressione sanguigna a valori normali quando le arterie si contraggono. Inoltre, con la diminuzione del flusso ematico, il metabolismo rallenta, riducendo così la richiesta di ossigeno da parte dei tessuti di tutto il corpo.

Presumibilmente, la contrazione delle arterie fa sì che il massimo rifornimento possibile di sangue, e quindi di ossigeno, sia a disposizione di quei tessuti che sono

Una foca di Weddell ritorna al foro praticato nel ghiaccio dagli studiosi, per respirare dopo un'immersione nell'oceano Antartico. Le immersioni a scopo alimentare durano in media 17 minuti e portano l'animale fino a una profondità di 500 metri o più; immersioni occasionali possono durare più di un'ora. Anche in brevi scorrerie, l'animale per sopravvivere deve rifornire di ossigeno i tessuti ed evitare la narcosi da azoto e l'embolia gassosa, provocate dall'elevata pressione. La fotografia è di Randall W. Davis del Sea World Research Institute di San Diego.

La distribuzione dell'ossigeno differisce molto nell'uomo e nella foca di Weddell. La foca immagazzina circa il doppio di ossigeno per chilogrammo di peso corporeo rispetto all'uomo e una frazione maggiore della sua riserva è contenuta nel sangue e nei muscoli, anziché nei polmoni.

I tessuti che continuano a essere irrorati normalmente dal sangue ossigenato (*scritte in rosso*) durante le immersioni in laboratorio comprendono quelli fondamentali per il controllo della direzione e del movimento, cioè la retina, il cervello e il midollo spinale. Anche le ghiandole surrenali, che producono un ormone che evita l'insorgere di convulsioni, ricevono un afflusso normale di sangue, come pure la placenta nelle foche gravide. Non sapendo quanto tempo rimarranno sott'acqua in un'immersione forzata, sembra che le foche si preparino al peggio riducendo drasticamente l'irrorazione sanguigna a molti altri tessuti, fra cui la muscolatura. Esaurita la loro scorta di ossigeno, i muscoli ottengono energia attivando un metabolismo anaerobico.

fondamentali per la sopravvivenza. I miei colleghi e io abbiamo recentemente misurato il flusso sanguigno in vari tessuti nella foca durante un'immersione in laboratorio. In accordo con i risultati precedenti, abbiamo trovato che la foca continuava a fornire sangue in quantità normale a retina, cervello e midollo spinale, tutti organi essenziali per il controllo della direzione e del movimento. (Come è logico attendersi, il cuore riceveva sangue in quantità ridotta come richiesto dal suo minore carico di lavoro.)

Altri due organi beneficiavano di un flusso sanguigno normale: le ghiandole surrenali e, nelle foche gravide, la placenta. Non è chiaro perché le ghiandole surrenali continuino a essere abbondantemente irrorate, ma il fatto che esse producano alte concentrazioni dell'ormone cortisolo può fornire un buon indizio. Vi sono prove che questo ormone stabilizzi le cellule nervose nel corso di un'immersione, impedendo così episodi convulsivi indotti dalla pressione elevata. È ovvio, invece, perché il sangue irrori la placenta: questo organo è indispensabile per la funzione respiratoria del feto e deve continuare a funzionare normalmente per consentire al piccolo di sopravvivere all'interno della madre.

I nostri studi sulla distribuzione dell'ossigeno hanno confermato anche che la foca blocca l'afflusso di sangue alla maggior parte degli altri sistemi e tessuti durante le immersioni in laboratorio. Quando l'irrorazione sanguigna cessa, molti dei tessuti coinvolti (come i reni) smettono di funzionare finché l'animale non riemerge per respirare, mentre altri, se devono compiere attività indispensabili, probabilmente continuano a funzionare con metabolismo anaerobico. Il prodotto secondario che rivela l'instaurarsi del metabolismo anaerobico è l'acido lattico; nelle foche che riemergono dopo un'immersione obbligata, il livello ematico di acido lattico è molto maggiore del livello di riposo.

Sebbene la possibilità di passaggio al metabolismo anaerobico possa essere importante quando manca l'ossigeno, esso può risultare estremamente pericoloso. Livelli elevati di acido lattico abbassano il *p*H del sangue e possono condurre ad acidosi, con insorgenza di crampi, ridotta contrattilità del cuore e anche morte. Le ricerche di laboratorio condotte da P. F. Scholander, della Scripps Institution of Oceanography, che negli anni trenta studiò il riflesso da immersione, hanno fatto supporre che la foca riesca a evitare l'acidosi limitando il metabolismo anaerobico ai muscoli scheletrici e agli altri tessuti che non vengono irrorati durante le immersioni. In assenza di flusso sanguigno, questi tessuti non possono riversare l'acido lattico nel sangue fino a quando l'animale riemerge; a quel punto, fegato, polmoni e altri organi hanno la possibilità di eliminarlo.

Gli esperimenti di laboratorio hanno anche tentato di spiegare come la foca di Weddell riesca a neutralizzare gli effetti della pressione esterna. Oltre ad avanzare l'ipotesi che elevati livelli di cortisolo possano prevenire gli episodi convulsivi, tali studi hanno dimostrato che la foca non possiede gli spazi pneumatici, che negli altri mammiferi possono creare problemi. Inoltre, alcuni adattamenti le consentono di evitare la narcosi da azoto e l'embolia gassosa. I suoi polmoni sono piccoli in proporzione alla massa corporea e quindi hanno una capacità ridotta di immagazzinare azoto, che potrebbe disciogliersi nel sangue durante l'immersione. Per di più, l'animale espira prima di immergersi: l'effetto più ovvio è quello di diminuire la galleggiabilità che ostacola la discesa, ma un ulteriore vantaggio è la riduzione del volume di gas presente nei polmoni.

Durante l'immersione, la pressione idrostatica che si esercita sulla flessibile gabbia toracica dell'animale espelle dagli alveoli polmonari la maggior parte dell'azoto rimasto, spingendolo nel sistema bronchiale. Studi anatomici effettuati da Gerald L. Kooyman e collaboratori alla Scripps Institution mostrano che i bronchi e i bronchioli della foca sono irrobustiti da anelli cartilaginei che consentono alle vie aeree di fungere da solido serbatoio per i gas. Poiché tali vie, a differenza degli alveoli, non sono a diretto contatto con il sangue, l'azoto ivi contenuto non viene liberato in circolo. Nell'uomo i bronchi e i bronchioli, sottoposti a un'elevata pressione, si chiuderebbero completamente e non potrebbero immagazzinare l'azoto in eccesso.

Parecchi anni fa, Kooyman determinò anche, con immersioni forzate in camera iperbarica, che i polmoni della foca collassano quando l'animale arriva a profondità comprese tra 50 e 70 metri. Il collasso dei polmoni arresterebbe la diffusione dell'azoto nel sangue, limitando la concentrazione ematica di tale gas.

Applicando al dorso di una foca un dispositivo costituito da un calcolatore che controlla un sistema per il prelievo automatico del sangue, Roger D. Hill, l'autore e i loro colleghi hanno misurato le risposte fisiologiche delle foche in immersioni libere nell'oceano. Nelle immersioni più brevi di 17 minuti, la risposta è stata meno drastica che nelle prove eseguite in laboratorio.

Più recentemente, Kooyman e collaboratori hanno effettuato alcune ricerche sul campo, che hanno indotto molti di noi a chiedersi se la foca in libertà reagisce all'immersione nel modo che avevamo osservato in laboratorio. Quando il gruppo di Kooyman tentò di studiare foche che si immergevano spontaneamente in mare, potè constatare che gli animali a volte non manifestano un vero riflesso da immersione. Quei primi studi, tuttavia, non furono definitivi perché l'attrezzatura disponibile non permetteva di seguire immersioni complete.

Solo dopo che Roger D. Hill del Massachusetts General Hospital ebbe messo a punto uno specifico programma ed ebbe costruito un calcolatore da otto bit funzionante a batteria, che rendeva possibile la valutazione delle risposte fisiologiche e metaboliche della foca nel corso di immersioni spontanee nell'oceano, il nostro gruppo, costituito da studiosi provenienti da laboratori di tutto il mondo, si riunì alla stazione di ricerca della National Science Foundation sulla costa del McMurdo Sound, nell'Antartide. Con un po' di fortuna, il calcolatore portatile, che pesava meno di due chilogrammi e possedeva 64 chilobytes di RAM (*random-access memory*), ci avrebbe consentito di chiarire fino a che punto la foca di Weddell manifestava il riflesso da immersione nel suo habitat naturale e in che modo riusciva a neutralizzare l'elevata pressione.

Il calcolatore di Hill, che era contenuto in un involucro a tenuta stagna in grado di resistere a profondità di 500 metri, svolgeva numerose funzioni. Per parecchi giorni, esso registrò il battito cardiaco alle varie profondità a intervalli di tempo prestabiliti. Inoltre controllava una pompa elettrica in grado di prelevare fino a 7 campioni di sangue arterioso a tempi (per esempio, ogni 10 minuti durante l'immersione) e a profondità determinati. Dopo il prelievo dei campioni, il calcolatore pompava il sangue in un sacchetto o in siringhe collegati a un cavetto in fibra ottica lungo circa due metri, che assicurava i collegamenti quando la foca riemergeva.

Prelevammo alcune foche da colonie vicine e le trasportammo in slitta fino al luogo prescelto per studiarle, un foro di un metro di diametro praticato attraverso uno strato di ghiaccio di due metri di spessore. Anestetizzammo un soggetto con la tecnica incruenta messa a punto da Robert C. Schneider del Massachusetts General Hospital, inserimmo i necessari cateteri e collegammo il calcolatore, che era fissato a un foglio di gomma incollato alla pelliccia del dorso dell'animale. (Con la muta estiva, l'apparecchiatura venne eliminata facilmente).

Equipaggiata con il nostro calcolatore e riavutasi dall'anestesia totale, la foca era libera di entrare nel foro e di allontanarsi. Eravamo fiduciosi che sarebbe tornata con i campioni di sangue e tutti

i dati previsti, perché avevamo adottato un metodo operativo abbastanza affidabile ideato da Kooyman. Sapendo che le foche di Weddell possono nuotare solo per pochi chilometri sott'acqua, egli praticò i fori necessari per studiare gli animali a grande distanza uno dall'altro nella spessa calotta di ghiaccio; in tal modo, le foche erano obbligate a ritornare al punto di partenza per respirare.

Dopo che un soggetto si era allontanato, allestivamo una baracca (con una grande apertura nel pavimento) sopra al foro praticato nella calotta, per fornire riparo sia a noi sia al calcolatore che in un secondo tempo avrebbe recuperato i dati dal calcolatore trasportato dalla foca. Quando l'animale ritornava, collegavamo rapidamente il cavetto di fibra ottica al calcolatore fisso. In meno di 10 secondi quest'ultimo raccoglieva i dati immagazzinati e, se necessario, impartiva nuove istruzioni.

Osservammo, come già Kooyman in precedenza, che circa il 95 per cento delle immersioni volontarie delle foche duravano meno di 20 minuti. Esse erano tendenzialmente immersioni per procurarsi il cibo, nelle quali cioè gli animali puntavano verso una preda e poi facevano ritorno. Immersioni di durata superiore a 20 o 30 minuti, ossia il rimanente 5 per cento, venivano intraprese solo quando le foche compivano esplorazioni a lungo raggio o dovevano sfuggire ai predatori.

La concentrazione di emoglobina (*a*), il pigmento trasportatore di ossigeno dei globuli rossi, aumenta durante i primi 10 - 12 minuti di un'immersione volontaria nell'oceano, impedendo al livello di ossigeno nel sangue della foca (*b*) di diminuire bruscamente. Nelle immersioni che durano più di 17 minuti (*linea tratteggiata*), in cui i muscoli attivano un metabolismo anaerobico, l'afflusso di emoglobina nel sangue equilibra efficacemente il consumo di ossigeno per circa 15 minuti, mentre in quelle più brevi (*linea continua*), in cui i muscoli evidentemente bruciano ossigeno, l'ulteriore apporto di emoglobina non può bilanciarne del tutto il consumo da parte dei tessuti; quindi il livello di ossigeno nel sangue diminuisce gradualmente fin dall'inizio.

La pressione parziale di azoto nel sangue cessa di aumentare quando la foca raggiunge la profondità di 40 metri, perché i polmoni, collassando, non liberano più il gas. Il collasso limita la quantità di azoto che può disciogliersi nel sangue durante un'immersione e contribuisce così a proteggere la foca dalla narcosi da azoto e dall'embolia gassosa. Con il procedere dell'immersione, il passaggio di azoto nei muscoli e nello strato adiposo riduce ulteriormente i rischi.

Studiando la distribuzione di ossigeno ai tessuti, Peter W. Hochachka dell'Università della British Columbia ha dimostrato che le foche non accumulano acido lattico nel flusso sanguigno durante o dopo immersioni che durino meno di 20 minuti. Ciò indica che durante le brevi immersioni abituali, la maggior parte di quelle compiute dalle foche, nei muscoli non si instaura il metabolismo anaerobico osservato nelle immersioni in laboratorio e l'irrorazione sanguigna deve essere continua. (Le masse muscolari dovrebbero produrre la maggior parte dell'acido lattico presente nel sangue, perché sono voluminose e svolgono lavoro durante le immersioni dell'animale.) Non dovendo eliminare acido lattico in seguito all'immersione, la foca è spesso in grado di ricominciare a pescare pochi minuti dopo essere tornata in superficie per respirare.

Ci siamo chiesti come la foca di Weddell rifornisca di ossigeno i muscoli senza privare di un adeguato rifornimento di questo gas il cervello e altri tessuti vitali. Nessuno ancora conosce la risposta, ma Hill ha potuto osservare che il battito cardiaco rallenta all'inizio di ogni immersione, ma non mantiene un ritmo costante durante le spedizioni più brevi; al contrario, esso accelera o rallenta a seconda della velocità di nuoto della foca, senza mai superare il ritmo di riposo. Quando il battito cardiaco accelera, anche la gittata cardiaca deve aumentare e questa quantità extra di sangue potrebbe benissimo essere diretta verso i muscoli scheletrici. Se così fosse, il blocco totale dell'irrorazione sanguigna ai muscoli, a lungo considerato un evento che si verifica regolarmente in ogni immersione, in realtà non avverrebbe nella maggior parte delle immersioni spontanee.

Contrariamente a quanto accade durante le immersioni a scopo alimentare, le occasionali escursioni prolungate della foca inducono la risposta osservata normalmente in laboratorio. Le lunghe scorrerie sono caratterizzate da accentuata bradicardia, con modesta variazione del ritmo cardiaco. Dopo (ma non durante) questo tipo di immersioni, la foca di Weddell libera acido lattico nel sangue, il che indica che l'animale esclude dall'irrorazione sanguigna i muscoli e conserva rigorosamente l'ossigeno durante l'immersione. Dopo aver attivato il metabolismo anaerobico a livello muscolare, la foca può restare sott'acqua per un'ora o più. Deve però pagare un prezzo: quando infine essa riemerge, non può immergersi di nuovo finché non ha eliminato l'acido lattico liberato dai muscoli, un processo che può durare anche un'ora. Perché immersioni anche brevi in laboratorio inducono la risposta tipica delle immersioni di lunga durata in mare? In laboratorio la foca non sa per quanto tempo dovrà stare immersa e così si prepara al peggio.

Altri studi sul rifornimento di ossigeno indicano che all'inizio di un'immersione, sia a scopo alimentare, sia esplorativa, la foca aumenta la concentrazione dei globuli rossi in circolo, massimizzando così il livello di emoglobina nel sangue e, quindi, la quantità di ossigeno disponibile per i tessuti. Jesper Qvist dello Herlev Hospital di Copenaghen, che ha fatto parte del nostro gruppo nell'Antartide, ha scoperto che la concentrazione di globuli rossi nelle arterie aumenta del 50 per cento nei primi 10 o 15 minuti di un'immersione. Contrariamente ai risultati di laboratorio ottenuti dal mio gruppo, che facevano ritenere che tali livelli fossero sempre elevati, Qvist ha dimostrato che i globuli rossi rappresentano inizialmente solo il 35 o il 40 per cento del volume del sangue circolante e poi salgono al 60 per cento nel corso dell'immersione (i livelli tornano alla normalità entro 10 minuti dalla riemersione).

Da dove potrebbe venire questo apporto di nuovi globuli rossi? Probabilmente dalla milza; è noto che quest'organo, peraltro poco conosciuto, si contrae quando il sistema nervoso simpatico è attivato, come si verifica quando un mammifero si immerge o è spaventato (la paura, perciò, può spiegare perché alcune foche mostrino elevati livelli di globuli rossi quando sono tenute in laboratorio). La contrazione della milza potrebbe immettere nel sistema venoso estremamente dilatabile della foca la sua riserva di globuli rossi ricchi di ossigeno; successivamente il cuore potrebbe distribuirli al circolo arterioso secondo necessità. Quando la foca torna in superficie per respirare, i globuli rossi circolanti verrebbero subito ricaricati di ossigeno e nuovamente immagazzinati.

La grossa milza della foca probabilmente accumula globuli rossi ricchi di ossigeno quando l'animale respira in superficie (*a*) e li immette in circolo quando si immerge (*b*). Il meccanismo potrebbe spiegare gli elevati livelli di emoglobina presenti nel sangue all'inizio di un'immersione nell'oceano. Secondo l'autore, la milza può immagazzinare circa 24 litri di globuli rossi. Essa si contrae quando inizia l'immersione, riversando gran parte del suo contenuto nella vena porta e poi, attraverso il fegato, nella vena cava posteriore, che è dilatabile. Uno sfintere permette a questo serbatoio di liberare i globuli rossi verso il cuore e nel circolo arterioso, quando è necessario.

Il ruolo di rifornitore di ossigeno attribuito alla milza non è nuovo; si sa che la milza di un cavallo da corsa mette in circolo globuli rossi pochi minuti dopo che l'animale ha iniziato un intenso allenamento. Una prova più diretta dell'importanza della milza per la foca viene dagli studi anatomici di parecchi mammiferi. Confrontando il peso dell'organo con il peso corporeo, ho trovato che le dimensioni della milza della foca di Weddell sono particolarmente grandi, uguagliate solo nell'elefante marino australe, un'altra specie in grado di compiere lunghe immersioni. Nell'uomo, nel cane e perfino nella balena, la milza è considerevolmente più piccola rispetto alla massa del corpo. Sulla base delle dimensioni dell'organo e del livello a cui arriva la concentrazione di emoglobina durante un'immersione, i miei colleghi e io stimiamo che la foca di Weddell immagazzini approssimativamente il 60 per cento della sua riserva totale di globuli rossi nella milza, mentre l'uomo ne conserva, in quella sede, meno del 10 per cento. In effetti, la milza della foca assomiglia a una specie di autorespiratore contrattile per la sua capacità di immagazzinare e rilasciare l'ossigeno necessario per un'immersione.

Gli effetti dell'incremento di globuli rossi ossigenati diventano particolarmente evidenti nelle lunghe immersioni. Mentre la concentrazione di globuli rossi aumenta, il contenuto di ossigeno del sangue rimane costante, il che indica che la quantità consumata dal cervello, dal cuore e da altri tessuti indispensabili è in qualche modo rimpiazzata. Il medesimo andamento costante non si riscontra invece nelle immersioni a scopo alimentare, dove i muscoli consumano ossigeno; qui i livelli di ossigeno cadono rapidamente. In questo caso, l'ossigeno bruciato dai muscoli della foca supera probabilmente la capacità della milza di immagazzinare sangue e, quindi, mettere in circolo globuli rossi.

Oltre a fornire ossigeno nel corso dell'immersione, l'afflusso di nuovi globuli rossi in circolo svolge probabilmente un altro ruolo importante: la diluizione dei gas disciolti nel sangue. Tale effetto spiegherebbe perché la concentrazione di anidride carbonica aumenti in misura sorprendentemente scarsa nel corso di un'immersione nell'oceano. La diluizione contribuirebbe a spiegare anche perché l'azoto non causi nella foca la narcosi o l'embolia gassosa. Gli studi sul campo indicano possibili soluzioni di altri quesiti. Per esempio, il nostro gruppo ha confermato la supposizione, già avanzata da altri, che i polmoni della foca collassino in immersione. Ciò si verifica intorno a 40 metri, leggermente prima di quanto fosse stato predetto.

Abbiamo determinato il punto di collasso basandoci sul lavoro di Konrad J. Falke dell'Università di Düsseldorf, che ha accuratamente misurato la pressione parziale di azoto in campioni di sangue prelevati dalle arterie di foche che si immergevano liberamente (la pressione parziale di azoto nel sangue è un buon indicatore della sua concentrazione, perché varia proporzionalmente a quest'ultima). Falke ha trovato che la pressione dell'azoto nel sangue, pari a 550 millimetri di mercurio quando l'animale respira in superficie, aumenta man mano che la foca scende in profondità, raggiungendo un massimo di 2000-2400 millimetri di mercurio quando l'animale arriva a 40 metri; oltre i 40 metri la pressione cala. Riteniamo che la caduta di pressione che segue al collasso dei polmoni avvenga non soltanto perché la milza mette in circolo nuovi globuli rossi, ma anche perché parte dell'azoto diffonde dal sangue verso i muscoli e lo strato adiposo.

Mentre studiavamo i molti e importanti adattamenti delle foche adulte a McMurdo Sound, eravamo sempre più incuriositi dalle risposte del feto della foca. Anch'esso mostra un riflesso da immersione quando la madre va sott'acqua? Il lavoro svolto a Fairbanks da Robert Elsner della Università dell'Alaska ha suggerito questa possibilità. Egli ha trovato che il battito cardiaco del feto,

analogamente a quello della madre, rallenta nel corso delle immersioni effettuate in laboratorio.

Non abbiamo ancora ottenuto una risposta soddisfacente alla nostra domanda, ma siamo riusciti a raccogliere alcuni dati sperimentali grazie a un apparecchio per il controllo del battito cardiaco collocato sul dorso di una foca gravida da Graham C. Liggins del National Women's Hospital di Auckland e da Hill. I dati raccolti dal calcolatore indicano che il battito cardiaco fetale rallenta davvero durante le immersioni libere, ma tale rallentamento è più graduale e meno marcato di quello della madre; inoltre il battito cardiaco accelera più gradualmente dopo che la madre è riemersa. Il feto «sa» quando la madre si immerge, sebbene non sia chiaro come riesca ad accorgersene. Dobbiamo completare altri studi volti a determinare se le alterazioni del ritmo cardiaco del feto siano accompagnate da cambiamenti nella gettata cardiaca e nella distribuzione del flusso sanguigno. Se questi cambiamenti avvengono effettivamente, significa che il feto immagazzina ossigeno per gli organi vitali quando la madre si immerge e ricostituisce le sue scorte quando essa riemerge.

I nostri studi sul campo dimostrano, nel complesso, che le risposte della foca di Weddell alle sue occasionali immersioni di lunga durata sono simili a quelle previste in base alle osservazioni in laboratorio. Il riflesso da immersione si verifica pienamente: il battito cardiaco rallenta e rimane basso per tutta la durata dell'immersione e i muscoli passano dal metabolismo aerobico a quello anaerobico, indicando così che sono esclusi dall'irrorazione sanguigna, probabilmente a causa della contrazione delle arterie.

L'andamento della maggior parte delle immersioni (quelle di minor durata) è invece piuttosto differente. Quando la foca intraprende un'esplorazione a scopo alimentare, la risposta all'immersione si modifica. Il ritmo cardiaco rallenta, ma è più variabile, e aumenta di nuovo quando la foca nuota più velocemente. Inoltre, i muscoli continuano a fare affidamento sul metabolismo aerobico; evidentemente, essi continuano a ricevere sangue, indicando che la contrazione vascolare può essere regolata. Sembra che la foca, all'inizio di un'immersione, «decida» se la sua scorreria sarà lunga o breve e se dovrà ricorrere o no a provvedimenti rigorosi per conservare l'ossigeno.

Il meccanismo tramite il quale la foca prende la decisione è uno dei problemi che devono ancora essere risolti. Grazie alla disponibilità di tecnologie sempre più progredite può non essere lontano il momento in cui si potrà rispondere alla domanda e raggiungere ulteriori conoscenze sui complessi adattamenti della foca di Weddell, uno dei più straordinari «sommergibili» del mondo.

La «Connection Machine»

La velocità di questo straordinario «calcolatore parallelo», con 65 536 unità di elaborazione in luogo di un'unica unità, potrebbe rivoluzionare nel prossimo futuro molte discipline e tra queste l'intelligenza artificiale

di W. Daniel Hillis

Negli ultimi 30 anni i calcolatori digitali hanno subito modificazioni notevoli. La potenza di calcolo che un tempo era contenuta in una stanza colma di valvole termoioniche ci è oggi offerta da dispositivi che stanno in una sola mano. Oggi si possono effettuare in pochi secondi calcoli anche molto complessi, la cui esecuzione un tempo avrebbe richiesto più giorni. Eppure, sotto certi rispetti fondamentali, l'architettura del calcolatore digitale è rimasta invariata dai tempi dell'ENIAC (una delle prime macchine digitali) fino all'attuale generazione di supercalcolatori. La maggior parte dei calcolatori moderni - dai supercalcolatori ai microcalcolatori - somigliano all'ENIAC poiché la memoria e l'unità centrale di elaborazione sono entità separate. Per eseguire un calcolo, i dati da trattare debbono essere estratti dalla memoria e portati nell'unità centrale; qui vengono elaborati e poi vengono riportati nella memoria.

Un'architettura di questo tipo si chiama sequenziale, perché le operazioni di elaborazione sono eseguite una alla volta. L'architettura sequenziale fu adottata soprattutto per ragioni di economia. Infatti ai primordi del calcolo digitale la memoria e l'unità centrale erano fatte di materiali diversi. Dato che la memoria era meno costosa dell'elaborazione, era opportuno rendere massima l'efficienza dell'unità di elaborazione a spese dell'efficienza della memoria: l'architettura sequenziale consente proprio questo. Oggi invece per fabbricare tanto la memoria quanto l'unità centrale vengono impiegate «fette» (*wafer*) di silicio con i circuiti incisi. Di norma in un calcolatore più del 90 per cento del silicio viene impiegato nella memoria. Mentre l'unità centrale viene mantenuta in uno stato di intensa attività, la gran parte del materiale se ne sta per lo più senza far nulla. Ma il silicio lavorato e incapsulato costa circa un milione di dollari al metro quadrato ed è quindi troppo prezioso per essere sprecato.

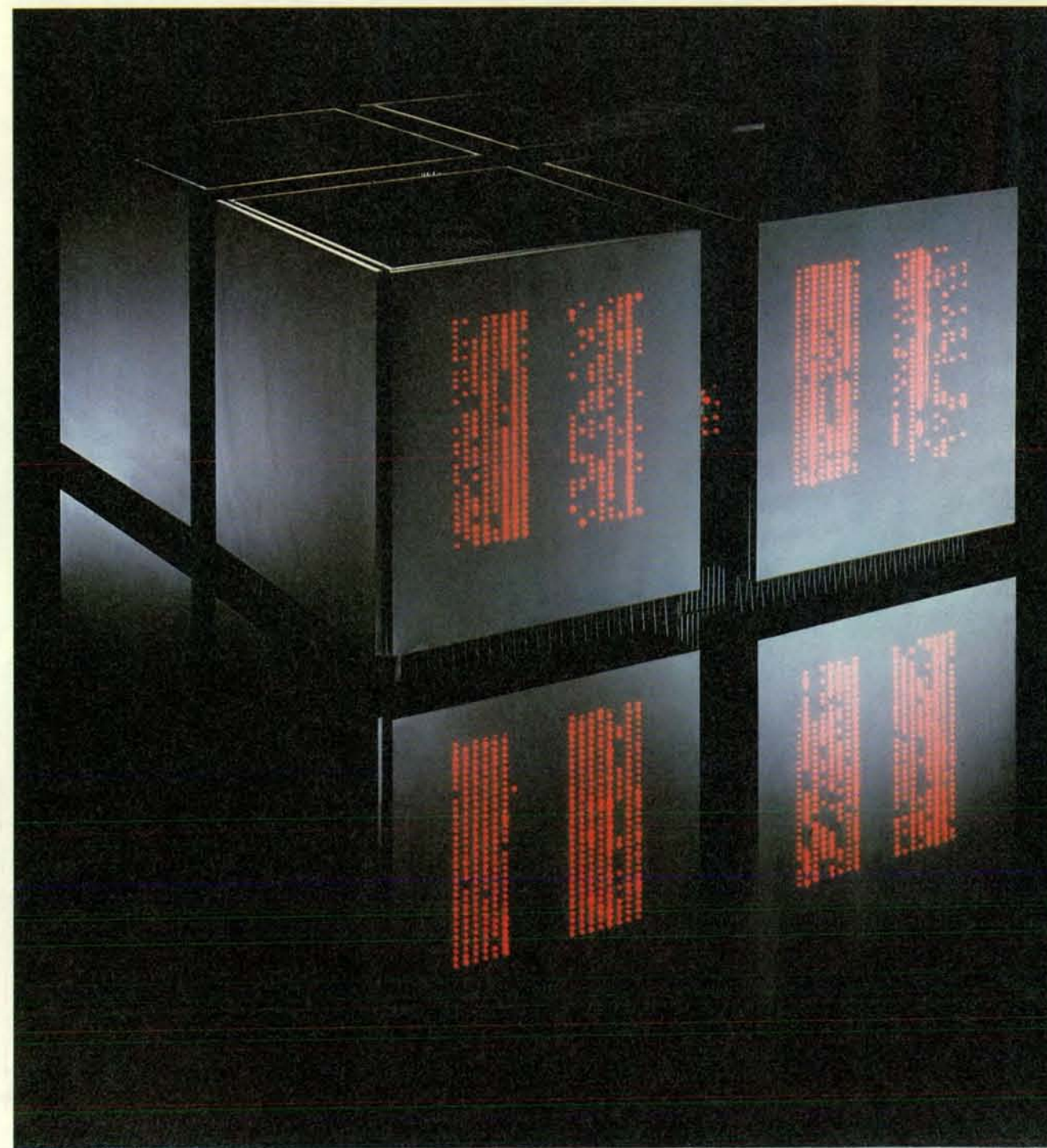
La soluzione di fondo di questo problema sta nel trovare un modo per unificare la capacità di elaborazione e la memoria. Ma come? Si potrebbero sfruttare molti piccoli elaboratori funzionanti simultaneamente, ciascuno fornito di una sua piccola memoria. Con questa architettura, che viene chiamata «a elaborazione parallela», sia la memoria, sia la capacità di elaborazione vengono utilizzate in modo molto efficiente. Ed è proprio questa l'impostazione che i miei colleghi e io abbiamo adottato nel costruire un calcolatore parallelo che abbiamo chiamato Connection Machine (in italiano «macchina connessionistica»). Questa macchina contiene 65 536 unità di elaborazione semplicissime. Benché ciascuna unità sia molto meno potente di un normale calcolatore personale, lavorando insieme esse riescono a eseguire parecchi miliardi di istruzioni al secondo. La macchina connessionistica è quindi uno dei calcolatori più veloci che siano mai stati costruiti.

Eppure l'aspetto più interessante di questa macchina non è tanto la sua velocità bruta quanto la sua flessibilità. Sono stati costruiti anche altri calcolatori paralleli, alcuni dei quali molto veloci per applicazioni particolari. Tuttavia, fuori del loro dominio specialistico queste macchine, come gli *idiots savants*, sono di solito molto impacciate. La macchina connessionistica, invece, può funzionare al massimo della sua velocità di elaborazione in un'ampia gamma di applicazioni. Ciò che le consente questa flessibilità, come spiegherò nelle pagine che seguono, è una rete di comunicazione, che permette alle numerose unità di elaborazione di scambiarsi informazioni secondo la configurazione più adatta al problema affrontato. La macchina connessionistica non è soltanto un prototipo: già ve ne è in commercio una decina di esemplari, che hanno cominciato a influire sul modo in cui vengono affrontati con l'elaborazione digitale certi problemi della fisica, dell'elaborazione delle immagini, del reperimento dei testi e anche dell'intelligenza artificiale.

Per comprendere i vantaggi dell'elaborazione parallela, è utile considerare la differenza che esiste nel modo di trattare la stessa immagine da parte di un calcolatore di tipo convenzionale e da parte del cervello umano. Dalle due immagini bidimensionali che si formano sulla retina, un essere umano è in grado - senza sforzo apparente - di ricostruire un modello tridimensionale del mondo e di mantenere integro tale modello anche quando le immagini bidimensionali si modificano rapidamente. Si può programmare un calcolatore in modo da fargli eseguire in parte questo compito, ma anche calcolatori velocissimi impiegano ore per fare ciò che il nostro cervello riesce a fare in qualche frazione di secondo (si veda l'articolo *La visione nell'uomo e nella macchina* di Tomaso Poggio in «Le Scienze» n.190, giugno 1984). Il cervello conserva il proprio vantaggio anche se i suoi componenti, i neuroni, sono milioni di volte più lenti dei transistori del calcolatore.

Perché dunque il cervello è tanto più veloce del calcolatore? La nostra comprensione dei circuiti cerebrali della visione non è perfetta, ma è chiaro che certe aree del cervello funzionano secondo i principi dell'elaborazione parallela: in queste zone l'immagine viene elaborata tutta in una volta. Il calcolatore invece esamina dell'immagine una minuscola porzione alla volta, come se la stesse osservando attraverso un piccolissimo buco di serratura. Nel calcolatore l'immagine è rappresentata sotto forma di una matrice di numeri, ciascuno dei quali corrisponde all'intensità luminosa di un punto particolare. Spesso questa matrice è un quadrato di 256 punti per lato. Dato che un calcolatore convenzionale opera su uno solo alla volta dei 65 536 punti del quadrato, un'operazione di elaborazione anche semplice comporta 65 536 passi.

La macchina connessionistica, invece, adibisce a ciascun punto dell'immagine un'unità di elaborazione distinta. Dato che ogni operazione può essere effettuata simultaneamente su tutti i punti, un calcolo che coinvolga l'immagine complessiva è rapido quanto un calcolo che riguardi un solo punto. Per esempio, per individuare tutti i punti dell'immagine più luminosi di un certo livello, una macchina sequenziale deve controllare uno dopo l'altro 65 536 minuscoli elementi, confrontando ciascuno di essi con il valore della soglia. Nella macchina connessionistica questi confronti vengono compiuti simultaneamente dalle 65 536 unità, ciascuna delle quali opera su un singolo elemento dell'immagine.

Il confronto con una soglia è particolarmente semplice perché può essere effettuato in modo indipendente da ciascuna unità di elaborazione. Nelle operazioni di calcolo più interessanti si richiede tuttavia che le unità di elaborazione si scambino informazioni via via che l'operazione procede. Consideriamo per esempio la convoluzione, una ope-

La *Connection Machine* («macchina connessionistica») è un cubo di 1,5 metri di lato, costituito da otto sottocubi. Ogni sottocubo contiene 16 pannelli disposti verticalmente, ciascuno dei quali porta 32 piastrine (*chip*). Ogni piastrina ha 16 unità di elaborazione, fornite di una piccola memoria. Le lucette rosse sui pannelli indicano lo stato delle piastrine e permettono di localizzare i guasti. Lavorando in parallelo, le 65 536 unità di elaborazione possono eseguire molti miliardi di istruzioni al secondo; questa macchina è uno dei calcolatori più veloci mai costruiti.

razione comune nell'elaborazione delle immagini. La convoluzione rende meno nitida l'immagine poiché fa la media tra ciascun punto e i suoi vicini immediati nella griglia bidimensionale. (La convoluzione, che è analoga a certe operazioni effettuate dal sistema visivo dell'uomo, serve a rimuovere piccoli particolari insignificanti e a dare risalto più netto agli oggetti significativi.)

Per effettuare la convoluzione, ciascuna unità di elaborazione deve leggere un valore dalle unità che contengono le informazioni relative ai punti situati a destra, a sinistra, sopra e sotto il punto in questione. Le unità devono insomma «parlare» tra loro. Un modo per attuare un tale schema di comunicazione sta nel collegare tra loro le unità in una griglia bidimensionale. Dato che ciascuna unità viene collegata alle sue quattro vicine immediate, questa griglia è in corrispondenza diretta con le vie di comunicazione necessarie alla convoluzione. In effetti alcuni calcolatori paralleli specializzati nell'elaborazione delle immagini hanno gli elementi collegati a formare una griglia bidimensionale; tuttavia questa configurazione va bene per la convoluzione, ma non per calcoli di altro tipo.

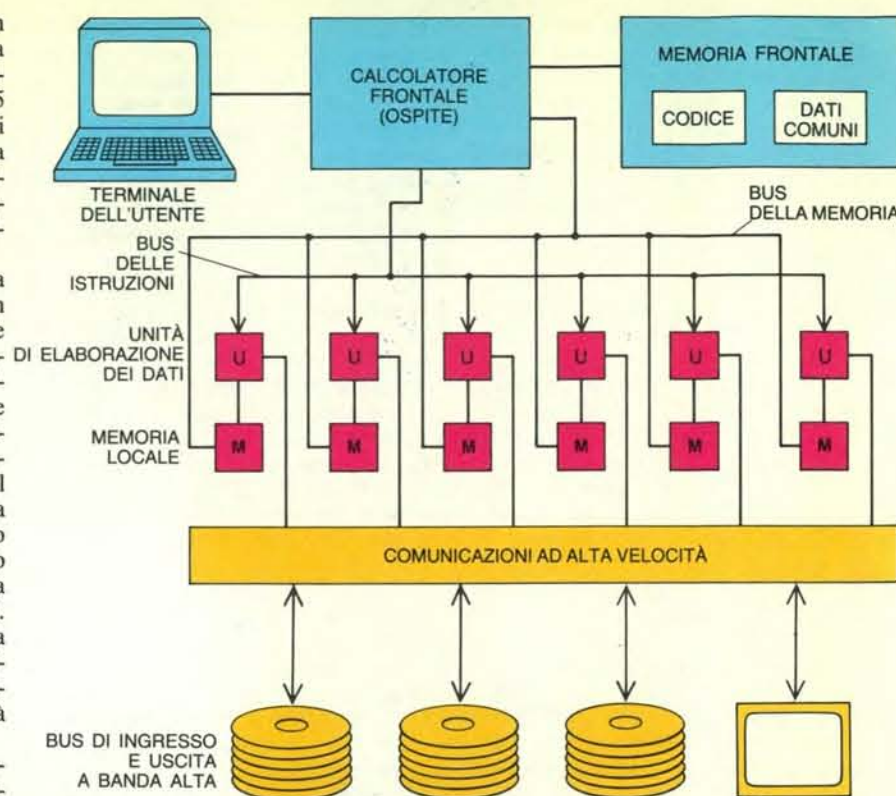
Per esempio per calcolare l'intensità media di tutti i punti dell'immagine, lo schema di collegamenti più conveniente assomiglia a un albero rovesciato. L'intensità media di un'immagine di 65 536 punti può essere calcolata ottenendo prima il valore medio di tutte le coppie di punti, poi il valore medio di tutte le coppie di coppie, e così di seguito. L'intensità media si può ottenere con 16 passi. Negli ultimi passi del calcolo sono richiesti scambi di informazione relativi a punti che nell'immagine sono molto distanti tra loro: perciò la griglia bidimensionale non è uno schema di collegamento conveniente.

I principi generali che si possono ricavare da questi esempi sono che ciascun tipo di calcolo può richiedere uno schema di collegamento distinto e che ciascuna unità di elaborazione può aver bisogno di comunicare con ogni altra unità. Quindi nel progettare la macchina connessionistica abbiamo scelto un tipo di rete di comunicazione che consente a ciascuna unità di comunicare con ogni altra. È grazie a questa flessibilità che il programmatore può scegliere l'algoritmo più adatto per risolvere il problema che gli interessa senza doversi preoccupare dei limiti imposti dallo schema dei collegamenti.

L'elemento fondamentale che si ripete nella macchina connessionistica è un circuito integrato composto di 16 piccole unità di elaborazione e di un dispositivo per l'instradamento delle comunicazioni. Ogni unità è dotata di 4096 bit di memoria. (Un calcolatore personale medio ha una memoria di 256 000 bit o più.) Le 16 unità sono incise su un'unica piastrina (*chip*) e le piastrine sono a loro volta incapsulate a gruppi di 32 su un unico pannello a circuiti stampati. La macchina contiene 128 di questi pannelli, che sono sistemati in un cubo di 1,5 metri di lato. Allo scopo di localizzare i guasti, ciascuna piastrina è collegata a una lucetta situata sul bordo del suo pannello; quando la macchina lavora, queste luci formano una figura variabile sulla faccia del cubo.

Le 16 unità di elaborazione di ciascuna piastrina sono collegate mediante un commutatore che consente di stabilire un collegamento diretto fra due elementi qualunque. Per ottenere un collegamento diretto fra due qualunque delle 65 536 unità del sistema, sarebbero necessari più di due miliardi di fili, un'impresa impossibile. Invece si è collegato il dispositivo d'instradamento di ciascuna piastrina con altri 12 dispositivi dello stesso tipo. Questi «instradatori» sono collegati tra loro secondo uno schema chiamato cubo booleano a *n* dimensioni. Questo cubo a *n* dimensioni, che è una generalizzazione del comune cubo tridimensionale, possiede eccellenti proprietà di rete di comunicazione fra le unità di elaborazione.

Una descrizione matematica completa del cubo a *n* dimensioni esula dall'ambito di questo articolo, ma il principio generale non è difficile da comprendere. Il comune cubo tridimensionale può essere pensato come un membro di una successione di «cubi» corrispondenti a dimensioni spaziali diverse. Per esempio, un segmento può essere considerato come un «monocubo» o cubo unidimensionale. Collegando gli estremi di due monocubi si ottiene un «bicubo» cioè un quadrato; congiungendo i vertici di due quadrati si ottiene un «tricubo», cioè la figura che normalmente chiamiamo cubo. Allo stesso modo, collegando i vertici di due tricubi si ottiene un quadricubo (*si veda l'illustrazione a pagina 66*). Questo procedimento può essere ripetuto un numero arbitrario di volte, e si può dimostrare agevolmente che un dodecacubo possiede 2^{12} (cioè 4096) vertici, tanti quante sono le piastrine della macchina connessionistica.

Questo cubo booleano a *n* dimensioni è una configurazione utile per parecchi motivi. In primo luogo nel dodecacubo nessuna unità di elaborazione dista più di 12 fili da qualsiasi altra, e questo facilita le comunicazioni nella rete. In secondo luogo, l'architettura del cubo a *n* dimensioni si armonizza bene con la logica binaria del calcolatore. In un calcolatore digitale infatti tutti i dati sono registrati sotto forma di sequenze di bit, ciascuno dei quali vale 0 oppure 1. Ebbene ogni cubo contenuto nel cubo a *n* dimensioni possiede due sottocubi, che possono essere contrassegnati rispettivamente con 0 e con 1. Di conseguenza ciascun punto del dodecacubo ha un indirizzo unico costituito da una sequenza di 12 cifre binarie. Il primo bit indica in quale dei due endecacubi del dodecacubo è contenuto

Lo schema del sistema mostra come la macchina connessionistica funziona in collaborazione con un calcolatore di tipo convenzionale, chiamato ospite. L'utente del sistema interagisce con l'ospite tramite un comune linguaggio, modificato per la programmazione parallela. Invece di effettuare le operazioni ripetitive una alla volta, tuttavia, l'ospite le affida alla macchina connessionistica, la quale esegue le operazioni in parallelo. I risultati si possono ottenere mediante svariati dispositivi di ingresso e di uscita tra i quali visualizzatori ad alta risoluzione.

il punto considerato; il secondo bit specifica di quale dei due decacubi si tratta; e così via fino alla determinazione di un unico punto.

Questi indirizzi binari possono essere usati per instradare i messaggi che si scambiano le 4096 piastrine della macchina connessionistica. In ogni messaggio che circola nel sistema è contenuto un indirizzo del genere. Ricevuto il messaggio, l'instradatore ne esamina l'indirizzo un bit alla volta e poi lo avvia all'instradatore successivo, verso la destinazione. A sua volta questo instradatore rileva il messaggio, ne esamina l'indirizzo e lo fa proseguire. Pertanto, affinché un messaggio giunga a destinazione non occorrono mai più di 12 passi.

Questa rete di comunicazione ha parecchie caratteristiche che ne accrescono la velocità e la flessibilità. Un'utile proprietà della configurazione a dodecacubo è che fra due unità di elaborazione arbitrarie vi sono molte vie di comunicazione altrettanto efficienti. Perciò se una di queste è già occupata da un messaggio in corso di trasmissione, l'instradatore ha la facoltà di scegliere un'altra via semplicemente elaborando in un ordine diverso i bit dell'indirizzo (*si veda l'illustrazione a pagina 67*).

Questo sistema di comunicazione possiede anche un altro genere di flessibilità intrinseca. In certi casi la rete di comunicazione si comporta più o meno come un centralino telefonico, istituendo un collegamento fra due unità di elaborazione in modo da farle comunicare in maniera continua ed esclusiva. Nei casi più complicati, invece, i messaggi sono così lunghi e il sistema è tanto sovraccarico che gli instradatori debbono comportarsi come gli uffici postali: mettono cioè in serbo pacchetti di informazioni che in seguito vengono avviati a destinazione. Queste decisioni vengono prese dai dispositivi d'instradamento in base ai fili liberi nel momento in cui si deve effettuare una trasmissione.

Tali proprietà consentono alla macchina connessionistica di predisporre numerosi schemi di comunicazione diversi, a seconda del particolare problema affrontato. Una caratteristica importante del sistema è che questi particolari non sono percepiti dall'utente, che non ha bisogno di saperne sul dodecacubo più di quanto l'utente telefonico medio debba saperne sulla commutazione digitale. Il programmatore interagisce con la macchina connessionistica tramite un calcolatore convenzionale, che si chiama ospite e che impiega un sistema operati-

vo e un linguaggio di programmazione normali. Le unità di elaborazione della macchina connessionistica sono collegate con l'ospite più o meno come lo sarebbe una comune unità di memoria.

In un certo senso, la macchina connessionistica è la memoria dell'ospite e questa relazione permette un'integrazione semplice fra il calcolo in parallelo e il software esistente. I programmi della macchina connessionistica presentano una somiglianza sorprendente con i programmi di tipo convenzionale. La differenza principale è che molte operazioni che di solito vengono effettuate con iterazioni sono sostituite da operazioni singole, corrispondenti all'attività simultanea di molte unità della macchina connessionistica; le vie di comunicazione necessarie sono predisposte automaticamente dall'hardware d'instradamento.

Va osservato che nell'hardware di questo sistema non si trova alcuna ricercatezza. Per conseguire semplicità e affidabilità, nel progettare la macchina connessionistica abbiamo scelto tecnologie ben collaudate. Le singole unità di elaborazione sono piuttosto lente rispetto ai calcolatori odierni più veloci. La piastrina adottata è costruita con metodi analoghi a quelli impiegati per fabbricare i calcolatori personali e le calcolatrici tascabili. Ma, lavorando insieme, le 65 536 unità conferiscono alla macchina una grandissima velocità: in molte applicazioni la macchina riesce a effettuare più di due miliardi di operazioni al secondo e nei casi più favorevoli questo numero supera i 10 miliardi, circa 1000 volte di più di un normale calcolatore di grandi dimensioni.

Inquadrando la velocità della macchina in un contesto lievemente diverso, consideriamo le operazioni in virgola mobile, che costituiscono un punto di riferimento per la potenza di calcolo nelle applicazioni scientifiche con grandi volumi di numeri. Una operazione in virgola mobile è una moltiplicazione o divisione di due numeri, che viene espressa con notazione scientifica (per esempio $1,5 \times 10^2$). Di norma un supercalcolatore riesce a effettuare poche centinaia di milioni di operazioni in virgola mobile al secondo; la macchina connessionistica riesce ad eseguirne circa 2500 milioni.

A quali campi è stata applicata questa notevolissima capacità di «macinare» numeri? Come si è già accennato, alcune delle prime applicazioni hanno riguardato il trattamento e l'elaborazione delle immagini. Altre applicazioni hanno sfruttato il parallelismo intrinseco di certi fenomeni fisici. In ingegneria, il problema di calcolare il flusso d'aria sopra l'ala di un aereo o il rotore di un elicottero costituisce un esempio di come la macchina connessionistica possa rispecchiare il parallelismo della natura.

In natura la configurazione globale del flusso è il risultato di una miriade d'interazioni fra le molecole dell'aria, che urtano fra loro e contro la superficie dell'ala in movimento. Il tecnico (al quale interessa il flusso complessivo più che le singole interazioni molecolari) impiega un modello macroscopico semplificato, consistente in un sistema di equazioni differenziali alle derivate parziali. Le equazioni tuttavia sono impostate in parallelo: esse riguardano infatti le variazioni della pressione in piccoli volumi d'aria e le loro interazioni si sommano per fornire il flusso complessivo. Poiché le equazioni sono in parallelo, la macchina connessionistica fornisce una soluzione rapida ed efficiente del sistema.

Con un calcolatore parallelo, tuttavia, si può anche andare oltre le equazioni e avvicinarsi di più alla comprensione della realtà fisica. Il comportamento macroscopico di un fluido è in grandissima misura indipendente dalle proprietà fisiche delle singole molecole. Inoltre il comportamento qualitativo del fluido non varia se il numero delle particelle subisce una forte riduzione. È quindi possibile riprodurre con grande accuratezza i flussi macroscopici esaminando gli urti fra poche decine di milioni di particelle semplici generalizzate.

Stephen Wolfram, che lavora presso il Center for Complex Systems Research dell'Università dell'Illinois, a Urbana-Champaign, e un mio collega, James Salem, hanno ottenuto con questa tecnica modelli del flusso dei fluidi su superfici complesse. La loro simulazione, che coinvolgeva solo poche decine di milioni di particelle, era abbastanza rozza, dato che ciascuna di esse poteva muoversi solo in sei direzioni con velocità intere. Il sistema, tuttavia, è in grado di imitare con precisione il flusso di un fluido.

Il modo più semplice e logico di eseguire il calcolo del flusso di un fluido sarebbe quello di assegnare a ciascuna particella la sua unità di elaborazione. Tuttavia una simulazione riguarda in genere circa otto milioni di «particelle» e la macchina connessionistica, per quanto grande, contiene soltanto 65 000 unità circa. La soluzione di questa, come di altre analoghe difficoltà, sta nel programmare ciascuna unità in modo che essa si comporti come se fosse una sequenza di unità di elaborazione diverse, ognuna delle quali tratta una sola particella alla volta. Anche in questo caso i particolari dell'allestimento sono invisibili al programmatore, il quale stabilisce semplicemente quante «unità di elaborazione virtuali» sono necessarie. Di tutto il resto si occupano l'hardware e il

Il cubo booleano a *n* dimensioni illustra la topologia della rete che collega le unità di elaborazione della macchina connessionistica. Il cubo booleano a *n* dimensioni è una generalizzazione del cubo comune. Si possono costruire cubi in molte dimensioni, ciascuno a partire da quello avente una dimensione in meno. Un punto può essere considerato un cubo a zero dimensioni, uno «zerocubo». Collegando due punti si ottiene un «monocubo», cioè un segmento. Collegando due «bicubi» (quadrati), si ha il cubo tridimensionale comune. Due tricubi possono essere collegati per i vertici (angoli) per dare un «quadricubo». Ripetendo il processo si ottiene un «dodecacubo» con 4096 vertici. Le 4096 piastrine della macchina connessionistica sono collegate secondo un dodecacubo.

Il cubo a *n* dimensioni offre vie di comunicazione alternative fra le piastrine. La figura mostra vie alternative in un tricubo, ma lo stesso principio vale per il dodecacubo della macchina connessionistica. A ogni vertice del cubo a *n* dimensioni (dove sono situate le piastrine) può essere assegnato un indirizzo unico, come illustrato di seguito. Un tricubo individua tre coppie di piani. Ogni piano può essere contrassegnato con 0 e con 1 e al vertice si assegna un indirizzo di tre cifre, ciascuna delle quali indica in quale piano della coppia corrispondente esso giace (*1*). I messaggi sono avviati dai dispositivi di instradamento collocati su ciascun vertice, che leggono gli indirizzi e li elaborano cifra per cifra. Qui un messaggio viene inviato da 000 a 111. L'instradatore legge la prima cifra e manda il messaggio al punto 100 (*2*). Qui viene letta la seconda cifra (*3*). In 110 viene letta la terza cifra e il messaggio è inviato a destinazione (*4*). Al momento di inviare il messaggio, tuttavia, il collegamento tra 000 e 100 potrebbe essere occupato (*5*). In tal caso l'instradatore semplicemente legge prima un'altra cifra dell'indirizzo, scegliendo un altro cammino (*6*). Dopo vengono lette la prima e la terza cifra (*7*, *8*) e il messaggio viene recapitato all'indirizzo giusto.

software. È ovvio che se ciascuna unità deve simularne 250 in sequenza, il calcolo richiede un tempo 250 volte superiore rispetto a quello che occorrerebbe se ogni particella disponesse della propria unità di elaborazione.

Molte applicazioni interessanti della macchina connessionistica non hanno carattere numerico. I colleghi Brewster Kahle, Craig Stanfill e David Waltz hanno sfruttato il parallelismo di questo calcolatore per reperire documenti in ampie raccolte di testi. Il loro sistema si basa sul principio che ogni unità di elaborazione può essere programmata in modo da confrontare un documento appartenente a una grande base di dati con un «campione di ricerca», un paragrafo scelto per la sua importanza. Compiuto il confronto, le unità si scambiano l'informazione e ordinano i documenti in base al grado di accordo che essi presentano con il campione di ricerca.

Confrontare due brani di prosa per valutarne l'accordo non è un compito facile. Limitarsi a contare le parole che compaiono in entrambi i campioni è inutile, perché il conteggio è inquinato da parole aventi scarso contenuto, come articoli e congiunzioni. Perciò il sistema di reperimento dei documenti estrae da ciascun campione le frasi che ne racchiudono il contenuto sfruttando un dizionario e certe regole grammaticali. Ciascuna unità viene caricata con un brano diverso compresso in questo modo e il campione di ricerca viene trasmesso a tutte le unità di elaborazione della rete.

Il procedimento di confronto è relativamente semplice e, poiché vengono controllati 65 536 documenti alla volta, l'intera base di dati può essere esaminata quasi all'istante. Ordinare i brani in base al grado di accordo che presentano con il campione di ricerca è un compito più difficile, poiché richiede uno schema di comunicazione complesso fra le unità di elaborazione. In parallelo, tuttavia, l'operazione può essere effettuata in circa 50 millisecondi. A questo punto alcuni dei documenti in testa alla lista sono presentati all'utente del sistema, che può scegliere fra essi un nuovo campione di ricerca. Poiché i confronti sono fatti tutti in una volta, la raccolta completa di testi può essere vagliata più volte in breve tempo, il che assicura che tutti i brani pertinenti saranno trovati.

Il programma per il reperimento dei documenti può funzionare senza bisogno di nulla che assomigli alla comprensione del contenuto dei brani; ciò richiederebbe una notevole conoscenza generale del mondo, conoscenza che non è stata inserita nel sistema di reperimento.

Un campo di ricerca appassionante, che concerne la macchina connessionistica, è la stesura di programmi che siano in grado di imitare certi aspetti della razionalità umana.

Senza sforzo apparente il cervello dell'uomo riesce a compiere tanto l'elaborazione delle immagini bidimensionali, che conduce alla formazione di un modello tridimensionale, quanto i comuni ragionamenti basati sul buon senso. Per esempio tutti i bambini sanno dedurre che, se lo si lascia andare, il vaso prediletto della mamma cadrà. Il bambino è in grado di inferire che il vaso assomiglia di più a un piatto o a un sasso che a un uccellino o a un palloncino, i quali non cadono. Questa inferenza può essere ef-

fettuata correttamente nonostante informazioni in apparenza contrarie, come per esempio il fatto che il vaso può essere sferico come il palloncino.

È evidente che per gli esseri umani la facilità e la precisione di queste inferenze aumentano via via che la loro conoscenza del mondo si accumula. Per i calcolatori di tipo convenzionale vale il contrario: via via che il numero dei concetti aumenta, aumenta ancora di più il numero delle relazioni possibili fra essi. Dato che un calcolatore sequenziale può esaminare queste relazioni solo una alla volta, il suo ritmo di lavoro si abbassa drasticamente all'aumentare del numero dei dati generali. È stata la lentezza dei calcolatori convenzionali nei ragionamenti di buon senso a generare uno degli stimoli che ci hanno portato a progettare la macchina connessionistica.

Verso la fine degli anni settanta, quando ero laureando al Massachusetts Institute of Technology (MIT) cominciai a interessarmi al problema di simulare con i calcolatori i ragionamenti basati sul buon senso. Mi parve che un modo per evitare di impantanarsi, come accadeva ai calcolatori di tipo convenzionale quando si chiedeva loro di compiere semplici deduzioni sui fatti quotidiani, fosse quello di costruire una macchina che, nell'affrontare i collegamenti possibili fra i concetti, potesse esaminarne più di uno alla volta. Ciò accadeva nel lontano 1978. Nel 1985 l'idea era giunta allo stadio di un prototipo reale grazie anche a una sovvenzione concessami dalla US Defense Advanced Research Projects Agency, che si offrì di acquistare la prima macchina. A quel tempo avevo già abbandonato il MIT e avevo contribuito alla fondazione della Thinking Machines Corporation, la società costituita per realizzare e commercializzare la macchina connessionistica.

Oggi che la macchina connessionistica è una realtà fisica, i ricercatori di intelligenza artificiale l'impiegano per risolvere i problemi relativi ai ragionamenti basati sul buon senso. L'idea base di questa impostazione è di assegnare un solo concetto fondamentale a ciascuna unità di elaborazione. I collegamenti fra le unità possono allora essere sfruttati per rappresentare relazioni multiple fra concetti semplici. Nell'esempio elementare fornito sopra, un'unità potrebbe rappresentare il concetto di «vaso», un'altra quello di «mamma» e una terza quello di «piace». I collegamenti fra queste tre unità conterrebbero compiutamente la nozione che «Alla mamma piace il suo vaso», mentre altri collegamenti potrebbero rappresentare la forma del vaso, la sua composizione e la sua storia. Quando viene il momento di stabilire che cosa accadrebbe se il vaso venisse lasciato cadere, si possono esaminare in parallelo i collegamenti pertinenti.

Dato che oggi sono in funzione una decina circa di macchine connessionistiche, non vi è dubbio che tra breve esisteranno per queste macchine molti programmi nuovi, che probabilmente si collocheranno per la maggior parte nelle quattro aree accennate sopra: elaborazione di immagini, simulazione di processi fisici, esplorazione di banche di dati e intelligenza artificiale. Una delle difficoltà più gravi che s'incontrano nell'uso della macchina connessionistica è di abituarsi a pensare in termini paralleli. I programmatori hanno oggi accumulato una considerevole esperienza nella programmazione delle macchine sequenziali, tanto che quest'attività è diventata per loro quasi una seconda natura. Imparare a redigere programmi per macchine parallele obbliga, invece, a pensare in termini completamente diversi.

Questa difficoltà sarà più grave per la macchina connessionistica che per altri tipi di macchine parallele. Non si deve infatti credere che la macchina connessionistica sia l'unica rappresentante di questo genere; anzi in questo momento molti progetti diversi di macchine ad architettura parallela si trovano in fase di attuazione più o meno avanzata. Generalizzando molto, questi progetti rientrano in due vaste categorie: le architetture a «grana grossa» e quelle a «grana fine». Le macchine a grana grossa presentano relativamente poche unità di elaborazione collegate tra loro, ciascuna con una potenza di calcolo piuttosto grande; nelle macchine a grana fine vi sono moltissime unità di scarsa potenza.

Queste due classi di calcolatori paralleli formano in realtà una gamma continua: a un'estremità vi è il calcolatore sequenziale comune, che possiede il numero minimo di unità di elaborazione, cioè una; all'altro estremo della gamma vi sono le architetture simili alla macchina connessionistica, che comprendono un grandissimo numero di piccole unità di elaborazione. Per quanto certe aziende e studiosi molto qualificati perseguano l'impostazione a grana grossa, io ritengo che alla lunga sarà l'architettura a grana fine a dimostrarsi più feconda.

Nello scrivere un programma per una macchina a grana grossa ci si può attenere a concetti molto simili a quelli seguiti nella programmazione dei calcolatori sequenziali; i problemi sorgono quando si cerca di coordinare i programmi. Nel redigere un programma per la macchina connessionistica ci si trova invece in un ambito del tutto diverso di problemi e di possibilità. Per sfruttare tutto il potenziale della macchina sarà necessaria una nuova concezione del calcolo, che noi programmatori abbiamo appena iniziato ad apprendere. Questa nuova metodologia sarà senza dubbio fruttuosa e insieme stimolante.

Alcuni di questi frutti potranno derivare dal fatto che la macchina connessionistica può essere ampliata fino ad acquisire una potenza di calcolo molto maggiore senza alcuna variazione fondamentale di architettura. La maggior parte delle applicazioni previste per questa macchina potrebbe trarre vantaggio da un calcolatore molto più grande delle versioni attuali della macchina connessionistica; per questo il calcolatore è stato progettato in modo da consentire un cospicuo aumento del numero di unità di elaborazione. La macchina connessionistica può essere ampliata aggiungendo semplicemente a una macchina già esistente altre unità di elaborazione, dispositivi di comunicazione e memoria.

Per fare un esempio limite di ampliamento, immaginiamo un calcolatore parallelo con un miliardo di unità di elaborazione. Questa macchina potrebbe certo presentare alcune delle caratteristiche della macchina connessionistica, anche se vi sarebbero senz'altro molti problemi nuovi da risolvere. Costruita con la tecnologia attuale, una macchina con un miliardo di unità sarebbe grande quanto un palazzo e costerebbe venti volte di più dei più grandi calcolatori esistenti. Potrebbe peraltro eseguire qualcosa come 100 milioni di milioni (10^{14}) di istruzioni al secondo, cioè parecchi ordini di grandezza in più rispetto alla potenza di calcolo combinata di tutti i supercalcolatori esistenti.

La costruzione di una macchina computazionale di tal fatta presenta problemi tecnici sostanziali, ma questi si possono risolvere: i problemi veri sono quelli di immaginare come si potrebbe usare una siffatta potenza. Certi problemi di ingegneria, comprese alcune estrapolazioni degli esempi ricordati sopra, potrebbero trarre vantaggio da questa capacità enorme, ma in un certo senso essi sono banali. Le applicazioni degne di una macchina da un miliardo di unità di elaborazione sono quelle che comportano una modifica radicale del modo in cui concepiamo il calcolo.

Un calcolatore parallelo con un miliardo di unità di elaborazione potrebbe costituire la base per un servizio computazionale pubblico analogo a quelli del gas e dell'elettricità oggi esistenti. Come una centrale termoelettrica genera elettricità che viene trasferita agli utenti, così questo enorme calcolatore parallelo potrebbe fornire potenza di calcolo alla dotazione di robot e di stazioni di lavoro di un'intera città. Questa visione, forse un po' avveniristica, non è affatto irrealistica e offre un'idea dell'ampiezza delle trasformazioni che il calcolatore ad architettura parallela potrebbe portare.

Il rotore di un elicottero genera un flusso d'aria complesso, che può essere facilmente simulato in parallelo sulla macchina connessionistica. Ciascuna unità di elaborazione fornisce un modello della circolazione in un certo strato d'aria (*piccole suddivisioni nell'immagine*). La circolazione dell'aria in ciascuna sezione influisce sull'aria di ogni altra sezione. Queste interazioni vengono calcolate in parallelo. Particolari come la distorsione ondulare sugli strati inferiori sono importanti per prevedere le forze che si esercitano sulle pale. La simulazione è stata eseguita da T. Alan Egolf, dell'United Technologies Research Center e da J. P. Massar, un collega dell'autore.

La grafica al calcolatore è uno dei campi più fecondi dei calcolatori paralleli. L'illustrazione è stata ottenuta con la tecnica detta del tracciamento dei raggi, in cui ogni unità della macchina connessionistica viene assegnata a un diverso pixel (elemento di immagine). Le unità di elaborazione tracciano raggi luminosi che rimbalzano fra oggetti immaginari, qui sfere di vetro e figure di occhi. Le traiettorie dei raggi determinano il colore definitivo di ciascun pixel. L'immagine è di Karl Sims, del Media Laboratory del Massachusetts Institute of Technology.

L'origine del programma statunitense di guerra biologica

Documenti governativi declassificati di recente rivelano come venne avviato il programma di ricerca sulle armi biologiche, che, segreto durante la seconda guerra mondiale, è oggi controverso argomento di dibattito

di Barton J. Bernstein

«Ma perché la distruzione degli insetti nocivi deve essere tenuta segreta?» chiedeva perplesso Franklin D. Roosevelt al suo consigliere speciale Wayne Coy il 14 luglio 1943, quando gli Stati Uniti si trovavano ormai coinvolti nella seconda guerra mondiale da un anno e mezzo. Il Presidente stava esaminando una richiesta di 405 000 dollari da parte del Department of Agriculture per finanziare ricerche su insetti infestanti e malattie delle piante. Egli sapeva soltanto che il Bureau of the Budget non era al corrente del progetto, e che il War Department aveva dato istruzione al Department of Agriculture di mantenerlo segreto.

Roosevelt chiese a Coy di indagare sulla misteriosa iniziativa. Due giorni più tardi Coy riferiva al Presidente che esistevano buoni motivi per tenere segreti i dettagli del progetto e suggeriva di chiedere maggiori informazioni su quella attività a un certo George W. Merck. Questi dirigeva un ufficio civile del War Department, il cui compito era la ricerca sulle armi biologiche.

In confronto al Progetto Manhattan da due miliardi di dollari, che portò alla produzione della bomba atomica, il programma statunitense per le armi biologiche durante la seconda guerra mondiale era una iniziativa di impegno modesto: il progetto occupava circa 4000 persone, compresi gli scienziati, e il suo costo totale era di circa 60 milioni di dollari. Forse non dovrebbe sorprendere, quindi, che il presidente Roosevelt non ricordasse che il Department of Agriculture collaborava con il Chemical Warfare Service dell'esercito nel programma di sviluppo sulle armi biologiche.

Né Roosevelt né il suo successore, Harry S Truman, furono mai posti davanti alla decisione di dare l'ordine per un attacco con armi biologiche; ma la questione non ebbe fine con la seconda guerra mondiale. Gli Stati Uniti hanno continuato a sviluppare il loro arsenale di armi biologiche per molti anni dopo la guerra, e l'interesse per il finanziamento di tale ricerca è rinato in tempi recenti: dopo aver toccato il minimo durante le amministrazioni Nixon, Ford e Carter, gli stanziamenti per progetti di armi biologiche sono risaliti a 60 milioni di dollari nel corso del 1986.

Lo scorso settembre si è svolto un congresso internazionale per rivedere la Convenzione del 1972, che mette al bando lo sviluppo, la produzione e il possesso di armi biologiche, comprese le tossine, a scopi offensivi. Sia gli Stati Uniti sia l'Unione Sovietica sono firmatari della Convenzione, ma gli Stati Uniti hanno accusato i sovietici di violazioni. In seguito a contatti informali, delegazioni scientifiche di entrambi i paesi si sono riunite a Ginevra in aprile per cercare di elaborare misure accettabili dalle due parti per la verifica dell'ottemperanza al trattato. Considerando le discusse accuse statunitensi, la riluttanza sovietica ai controlli e le contestazioni reciproche sulle attività in questo campo, l'accordo del 1972 potrebbe essere in pericolo.

Queste circostanze conferiscono interesse sia allo studio delle origini del programma sia all'esame delle decisioni che impedirono l'impiego di armi biologiche da parte degli Stati Uniti nella seconda guerra mondiale. Con un centinaio di documenti, selezionati tra migliaia di rapporti statunitensi ottenuti su speciale richiesta e tra quelli degli archivi britannici un tempo segreti, ho potuto ricostruire le discussioni dell'epoca.

Nella prima guerra mondiale, agenti chimici come il cloro e l'iprite avevano ucciso o menomato più di un milione di soldati e civili. La reazione inorridita dell'opinione pubblica di fronte a queste morti spinse quaranta nazioni a firmare, nel 1925, il Protocollo di Ginevra, con il quale veniva proibito il primo impiego di armi chimiche e biologiche, senza però porre limiti alla ricerca, alla produzione e alla creazione di scorte di queste armi. Negli anni successivi, la maggior parte delle potenze industriali portò avanti attivamente programmi di sviluppo. Sebbene gli Stati Uniti avessero firmato il Protocollo di Ginevra, non lo ratificarono fino al 1975.

L'esercito statunitense iniziò a condurre ricerche sulle armi biologiche nel 1941 attraverso il Chemical Warfare Service, ma questi sforzi si concretizzarono solo nel 1942. Nel febbraio di quell'anno un comitato speciale istituito dalla National Academy of Sciences sottopose al Segretario alla guerra Henry L. Stimson un rapporto che conteneva raccomandazioni sul futuro del programma sulle armi biologiche. Stimson aveva richiesto il rapporto pochi mesi prima del bombardamento di Pearl Harbor.

Il comitato, composto da eminenti biologi come Edwin B. Fred dell'Università del Wisconsin e Stanhope Bayne-Jones della Yale University, concluse che un attacco nemico con armi biologiche avrebbe potuto provocare gravi danni alle persone, ai raccolti e agli animali. Sebbene il rapporto ponesse l'accento sulla difesa, richiedendo che si cominciassero a preparare vaccini e a proteggere le riserve d'acqua, il comitato raccomandò anche, in maniera piuttosto vaga, che gli Stati Uniti conducessero ricerche sul potenziale offensivo delle armi batteriologiche.

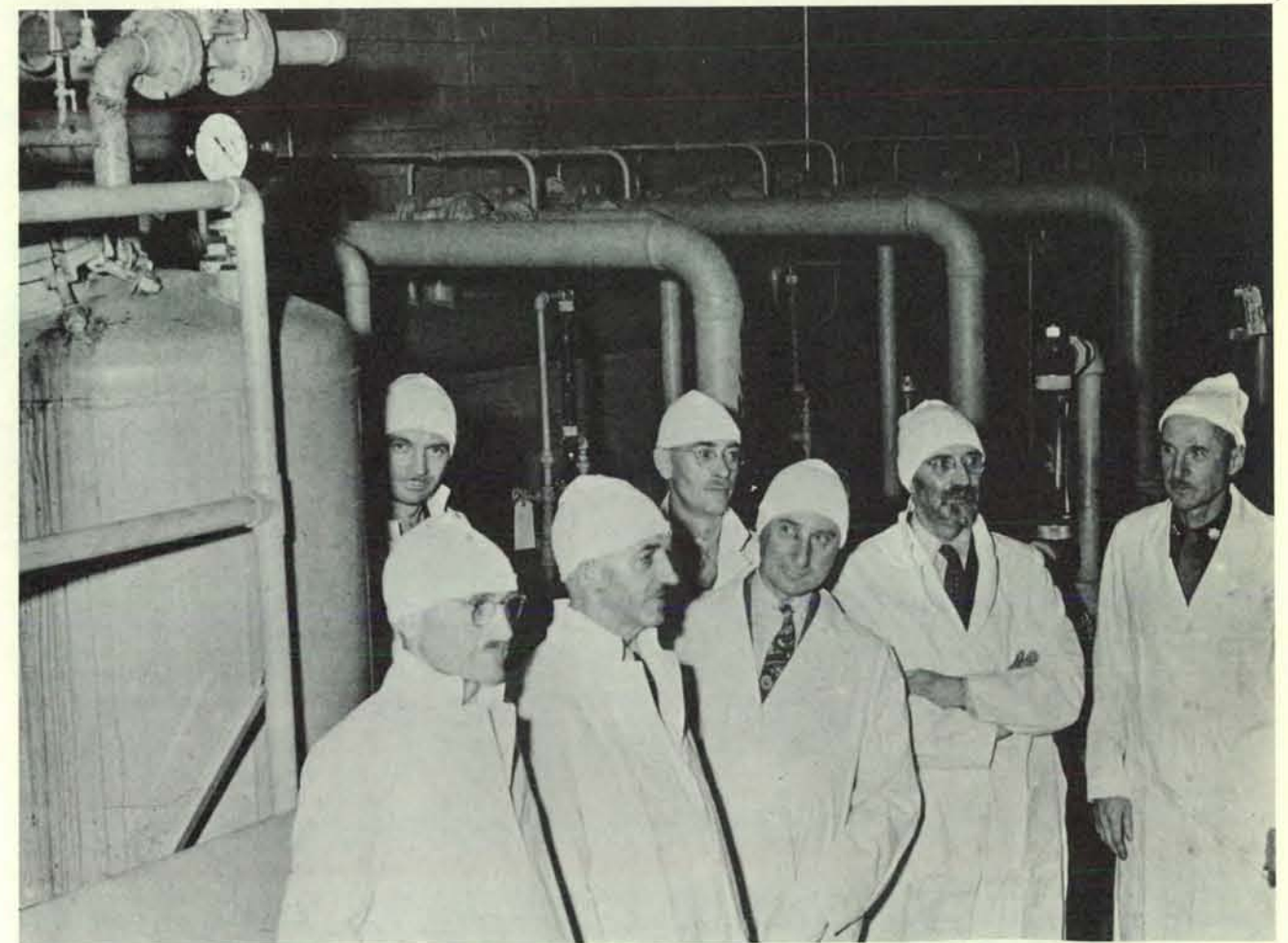
Spinto dagli avvertimenti degli scienziati, Stimson cercò di ottenere l'approvazione del Presidente per un programma ufficiale per lo sviluppo di armi biologiche che prevedesse la costituzione di un piccolo gruppo di consiglieri allo scopo di coordinare e dirigere tutte le ricerche governative. «Dobbiamo essere preparati» Stimson scrisse a Roosevelt in un memorandum dell'aprile 1942. «E il problema deve essere trattato con molta segretezza e insieme con grande energia.»

Stimson omise di menzionare che il Chemical Warfare Service aveva già iniziato ricerche sulle armi biologiche e il Presidente probabilmente non era informato del programma. Tuttavia il Servizio ricevette più tardi stanziamenti per milioni di dollari dai fondi dell'esercito e divenne determinante nel programma per lo sviluppo di armi biologiche più di quanto lo fosse il gruppo di consiglieri che lo dirigeva. Perché allora Stimson fece pressioni per formare il gruppo?

Probabilmente perché, come egli disse a Roosevelt, «la guerra biologica è una faccenda sporca». Nominando dei civili come controllori, Stimson sperava di legittimare le ricerche del Chemical Warfare Service. Mentre alcuni membri del comitato della National Academy of Sciences ritenevano che il programma dovesse essere gestito dal War Department, gli ufficiali superiori dell'esercito preferivano che l'incarico venisse affidato a un'agenzia di civili legata ai servizi militari. Stimson spiegò a Roosevelt il loro modo di ragionare: «Affidare la questione a una agenzia di civili ci aiuterebbe a evitare preoccupazioni eccessive nell'opinione pubblica al pensiero che il War Department stia esaminando la possibilità dell'impiego offensivo di queste armi».

Stimson suggerì di celare il gruppo di consiglieri per la «guerra dei germi» in una agenzia di servizi sociali, la Federal Security Agency, che sovrintendeva al Public Health Service e alla Social Security. Per dirigere il programma, egli voleva una personalità accademica che avesse familiarità con la ricerca universitaria e fosse esperto in amministrazione. Dopo la riunione di gabinetto del 15 maggio, Roosevelt ammise di non avere ancora letto il progetto del Segretario, ma gli disse di procedere comunque. Una settimana più tardi, Stimson discusse le proprie idee con il Segretario all'agricoltura Claude R. Wickard, il cui ente in seguito avrebbe preso parte alla ricerca coordinata dal gruppo di consiglieri, e con Paul V. McNutt, che dirigeva la Federal Security Agency.

All'inizio dell'estate, tre candidati avevano rifiutato l'offerta di dirigere il nuovo gruppo: l'economista Walter W. Stewart, che presiedeva la Rockefeller Foundation, il geografo Isaiah Bowman presidente della Johns Hopkins University e l'economista Edmund Ezra Day, presidente della Cornell University. Alla fine, in agosto, accettò il chimico George W. Merck, presidente della ditta farmaceutica Merck & Co., Inc.

Il gruppo ebbe la denominazione innocua di War Research Service (WRS), e cominciò a lavorare a metà del 1942 con uno stanziamento iniziale di 200 000 dollari. Estesi contatti con i maggiori biologi e medici permisero agli otto membri della commissione di avviare ricerche segrete in 28 università statunitensi, fra cui la Harvard University, la Columbia University, la Cornell University, l'Università di Chicago, la Northwestern University, la Ohio State University, l'Università di Notre Dame, l'Università del Wisconsin, la Stanford University e l'Università della California. Nel gennaio 1943, il WRS aveva incaricato William A.

Alcuni scienziati visitano le installazioni di Camp Detrick, che nel 1943 era il centro delle ricerche sulle armi biologiche del Chemical Warfare Service. I personaggi nella fotografia sono *(in primo piano, da sinistra)* N. Paul Hudson della Ohio State University, Guilford B. Reed della Queen's University, Charles A. Mitchell del Dominion Department of Agriculture, Everitt G. D. Murray della McGill University e il colonnello Oram C. Woolpert; in secondo piano sono James Craigie *(a destra)* dell'Università di Toronto e il colonnello Arvo T. Thompson.

The value of biological warfare will be a debatable question until it has been clearly proven or disproven by experience. Such experience may be forthcoming. The wise assumption is that any method which appears to offer advantages to a nation at war will be vigorously employed by that nation. There is but one logical course to pursue, namely to study the possibilities of such warfare from every angle, make every preparation for reducing its effectiveness and thereby reduce the likelihood of its use. In order to plan such preparation, it is advantageous to take the point of view of the agressor and to give careful attention to the characteristics which a biologic offensive might have.

DECLASSIFIED
E.O. 12356, Sec. 3.3

La fattibilità delle ricerche sulle armi biologiche viene sottolineata in questo estratto dalle raccomandazioni contenute in un rapporto del 1942 della National Academy of Sciences. Probabilmente il rapporto fu determinante nel convincere il presidente Roosevelt a creare il War Research Service. I timbri impressi sul documento indicano il passaggio di classificazione da segreto a pubblico. Il testo inglese dice: «L'utilità della guerra biologica rimarrà dubbia fino a quando non sarà stata chiaramente dimostrata o negata dall'esperienza. Quel momento potrebbe essere vicino. È saggio presumere che una nazione in guerra debba impiegare con estrema decisione qualsiasi sistema possa avvantaggiarla. Esiste una sola linea logica da seguire, vale a dire studiare le possibilità di impiego di queste armi da ogni angolazione, trovare ogni rimedio per ridurne l'efficacia e, di conseguenza, anche la probabilità che vengano usate. Per predisporre tali rimedi, è vantaggioso assumere il punto di vista dell'aggressore e considerare attentamente le possibili caratteristiche di un'offensiva condotta con armi biologiche.»

Hagan della Cornell University di studiare l'uso offensivo del botulismo e J.Howard Mueller della Harvard Medical School quello del carbonchio.

Il carbonchio e il botulismo rimasero i punti focali delle ricerche sulle armi biologiche nel periodo bellico. Si tratta di malattie mortali provocate da batteri resistenti e prolifici, con periodi di incubazione assai brevi, di pochi giorni o anche poche ore. Le spore del carbonchio, resistenti e virulente, possono essere inalate o assorbite attraverso tagli nella pelle; il botulismo invece è causato dall'ingestione della tossina botulinica.

Il WRS diede al Chemical Warfare Service la possibilità di ampliare largamente la sua attività nel campo delle armi biologiche, consentendogli di operare al di fuori delle strutture universitarie. Negli anni 1942 e 1943, il Servizio ricevette milioni di dollari per costruire laboratori di ricerca, il più importante dei quali fu Camp Detrick a Frederick nel Maryland (ora Fort Detrick), che costò quasi 13 milioni di dollari. Il Servizio assunse anche molti scienziati per lavorare sia qui, sia altrove.

Provenienti principalmente da facoltà universitarie, essi vinsero la loro ripugnanza verso lo sviluppo di strumenti di morte, perché tale lavoro sembrava necessario nella situazione eccezionale della seconda guerra mondiale. Già ai primi del 1942, Theodor Rosebury, microbiologo della Columbia University, affermò che «la probabilità che armi batteriologiche vengano usate contro di noi aumenterà sicuramente se un nemico sospetta che siamo impreparati a fronteggiare il pericolo e a ribattere colpo su colpo». Poco dopo, Rosebury entrò a far parte del laboratorio del Chemical Warfare Service e divenne uno dei responsabili delle ricerche a Camp Detrick. Egli scrisse più tardi : «Stavamo combattendo contro un incendio [l'Asse] e ci sembrò necessario correre il rischio di sporcarci le mani oltre che di bruciarcele».

Stimson e McNutt avrebbero potuto benissimo approvare questo modo di pensare, ma sarebbero rimasti interdetti nel venire a conoscenza dell'opinione di Rosebury su chi dovesse prendere le decisioni. Rosebury riteneva che le preoccupazioni morali degli scienziati del suo laboratorio dovessero governare l'impiego delle armi che essi stavano creando. Anni dopo, scriveva: «I civili, con o senza uniforme, hanno preso tutte le decisioni importanti; i militari di professione se ne sono tenuti fuori. Noi abbiamo risolto il problema etico proprio come altri uomini ugualmente buoni hanno risolto lo stesso problema a Oak Ridge, ad Hanford, a Chicago, a Los Alamos.»

In realtà, le cose andarono diversamente. Anche se non fu il Presidente in persona a stabilire la linea di azione del WRS, sembra chiaro che le decisioni importanti furono prese a Washington e non nel laboratorio.

Benché il suo interesse fosse rivolto in primo luogo ai servizi sociali e assistenziali, Paul McNutt non trascurò di occuparsi del programma segreto celato nella sua agenzia. Nel febbraio del 1943, McNutt informò il presidente Roosevelt che il fondo di 200 000 dollari del WRS stava per esaurirsi. Disse che il Presidente avrebbe dovuto decidere se «approfondire due o tre... progetti attualmente in corso.». In aprile, con l'approvazione di Stimson, McNutt richiese altri 25 000 dollari per l'anno fiscale 1943 e un totale di 350 000 dollari per il 1944. Due giorni più tardi, Roosevelt scrisse in calce alla richiesta di McNutt la laconica approvazione: «O.K. F.D.R.». Il finanziamento del 1944 per il WRS fu ancora aumentato alcuni mesi dopo, allorché Roosevelt lo portò a 460 000 dollari.

In accordo con l'esigenza di assoluta riservatezza del programma, McNutt non mise per iscritto particolari progetti o loro dettagli, persino nella sua corrispondenza con il Presidente. L'archivio di Roosevelt contiene meno di una decina di lettere e memorandum a proposito di armi biologiche. Dei pochi documenti del 1942 e 1943, la maggior parte tratta dei piccoli stanziamenti e delle disposizioni amministrative per il WRS. Probabilmente Roosevelt veniva informato sui milioni di dollari che venivano aggiunti agli stanziamenti per l'attività nel campo delle armi biologiche nelle discussioni con McNutt e Stimson o in incontri con il generale George C. Marshall, l'autorevole Capo di stato maggiore dell'esercito. Tuttavia nessuno dei documenti disponibili rivela che Roosevelt abbia ricevuto tali informazioni.

Nel frattempo il Chemical Warfare Service stava ampliando le sue strutture per sviluppo, sperimentazione e produzione. In aggiunta ai circa 2 chilometri quadrati di Camp Detrick, fu realizzata un'installazione di circa 8 chilometri quadrati, per prove sperimentali, sulla Horn Island a Pascagoula, nel Missouri. Un territorio di circa 650 chilometri quadrati, vicino al poligono sperimentale di Dugway nello Utah, fu destinato a esperimenti di bombardamento, e quasi 25 chilometri quadrati furono messi a disposizione per la costruzione di un impianto di produzione vicino a Terre Haute, nell'Indiana.

Anche la tecnologia progrediva. Con l'assistenza tecnica inglese, il Servizio fece considerevoli progressi nella produzione di bombe biologiche, e nella seconda metà del 1943 cominciò a lavorare su bombe al carbonchio da 225 chilogrammi. Queste bombe contenevano 106 «bombette» da circa 2 chilogrammi che dovevano disperdersi e aprirsi nell'impatto. Tali bombe non erano state sperimentate, ma era nota la quasi inevitabile mortalità del carbonchio.

Il Servizio riuscì anche a produrre il botulino, una delle più potenti tossine gastrointestinali. Anche il semplice assaggio di un cibo contaminato dalla tossina è di solito sufficiente a causare una grave intossicazione o la morte. Nelle epidemie naturali il tasso di mortalità è compreso tra 16 e 82 per cento, ma attraverso modifiche alla tossina e al meccanismo di liberazione, gli scienziati di Camp Detrick speravano di produrre un'arma quasi sicuramente mortale.

Incoraggiato dai progressi compiuti, all'inizio del 1944 il Chemical Warfare Service cominciò a fare pressioni per uno stanziamento aggiuntivo di due milioni e mezzo di dollari per produrre bombe al carbonchio e al botulino. In questo modo, il Servizio avrebbe potuto produrre in un mese 275 000 bombe al botulino oppure un milione di bombe al carbonchio, ma occorreva tempo per costruire le fabbriche. Pertanto, le armi non sarebbero state disponibili in quantità significative fino al 1945, quando, secondo le previsioni degli strateghi militari, solo il Giappone sarebbe rimasto in guerra.

Il Servizio ottenne i fondi richiesti. Sebbene l'idea di un arsenale biologico con cui fronteggiare i giapponesi potesse sembrare allettante, una minaccia più immediata può aver contribuito ad accrescere l'importanza delle ricerche condotte dal Servizio. Nei primi mesi del 1944, i servizi di spionaggio dei paesi alleati cominciarono a temere che le nuove potenti «bombe volanti» V-1 tedesche avrebbero potuto presto essere dirette contro la Gran Bretagna o le truppe in Normandia, e che le testate dei missili avrebbero potuto essere caricate con agenti per la guerra batteriologica. Gli esperti avvertirono che l'alto comando tedesco stava fronteggiando una crisi strategica; esso stava riunendo tutte le proprie risorse e avrebbe potuto fare uso di armi biologiche per conseguire un vantaggio decisivo.

Le analisi erano basate sulla cosiddetta ipotesi del male peggiore e non erano certo confortanti; nel giugno del 1944, gli Stati Uniti avevano potuto approntare, al più, poche bombe al carbonchio a titolo sperimentale. Certamente nessuna bomba era disponibile per l'impiego contro un nemico.

Per scoraggiare la Germania dal dare inizio a un attacco biologico, i comandi militari dettero disposizioni di vaccinare circa 100 000 soldati contro il botulino, sperando di convincere i tedeschi che le truppe alleate stavano preparandosi alla controffensiva biologica. Nel caso che la Germania avesse messo davvero in atto un attacco biologico, con tutta probabilità le forze angloamericane avrebbero risposto con i gas.

La Germania non si comportò come era stato previsto, ma usò solamente esplosivi convenzionali nelle V-1. In realtà, per ragioni che rimangono ancora ignote, Hitler aveva proibito le ricerche su armi biologiche a scopo offensivo. Il programma americano, sviluppato sostanzialmente per fronteggiare una minaccia tedesca che non era mai esistita, rimase senza applicazioni.

A Camp Detrick, il lavoro procedeva rapidamente. Nel maggio del 1944, Stimson e McNutt presentarono a Roosevelt un breve riassunto della ricerca, che dedicava solamente cinque righe agli sviluppi scientifici. Si sarebbe però potuto dire molto di più. Attraverso il Chemical Warfare Service, un impianto aveva ricevuto l'autorizzazione a produrre un milione di bombe al carbonchio, mentre si stavano facendo progressi nelle tecniche di dispersione a corto raggio del botulino sotto forma di pasta. Nel novembre dello stesso anno, Merck mandò un rapporto a Stimson e Marshall - ma non a Roosevelt - che menzionava in maniera enigmatica ricerche su altri quattro «agenti anti-uomo». Da altre fonti di informazione si può ipotizzare che questi agenti fossero la brucellosi (o febbre maltese), la psittacosi, la tularemia e la morva (o cimurro).

Merck sosteneva che il Chemical Warfare Service stava anche sviluppando «almeno cinque agenti per l'impiego contro le piante.» (Si tratta in realtà di prodotti chimici, ma all'epoca venivano considerati parte del programma biologico perché potevano distruggere i raccolti.) Un sesto composto, il tiocianato di ammonio, era previsto per la distruzione dei «giardini giapponesi».

Questi sviluppi sono contenuti in do-

La *Black Maria* era la costruzione, coperta di carta catramata scura, che nel 1944 ospitava gli esperimenti sulle armi biologiche a Camp Detrick. Un soldato armato di fucile mitragliatore occupava la torre di guardia sulla sinistra. Il fabbricato venne distrutto dopo la guerra.

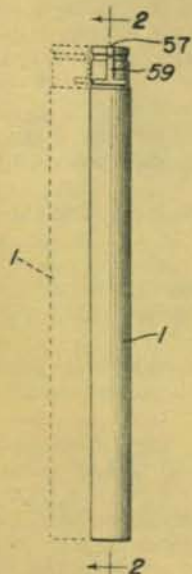
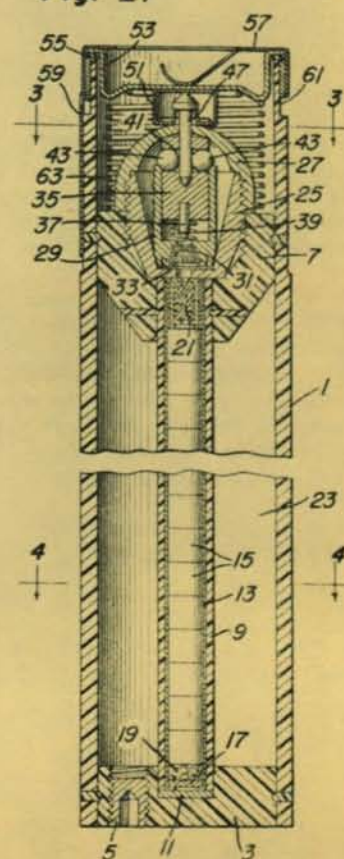
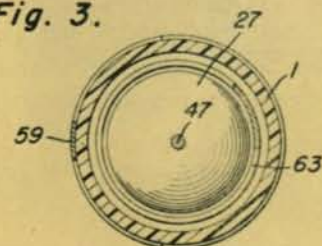
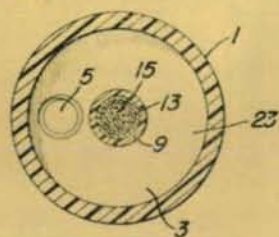
dici righe del breve rapporto di Merck sulle armi biologiche presentato a novembre. Il documento si trova fra le carte dell'archivio di Stimson a Washington, ora rese pubbliche. Non c'è alcuna prova che il Segretario o il Presidente abbiano prestato alcuna attenzione agli aspetti scientifici del programma.

Roosevelt trascurò non solo i lati scientifici ma anche quelli politici della guerra biologica. Malgrado il considerevole progredire dei lavori a Camp Detrick e il timore di una offensiva biologica tedesca, il Presidente sembrava aver pensato ben poco a questo problema. Nel 1942 e nuovamente nel 1943, Roosevelt si era pubblicamente impegnato a non ordinare attacchi con i gas, ma aveva minacciato rappresaglie nel caso l'Asse li avesse usati. Sembra che non abbia mai considerato l'idea di fare una simile dichiarazione per le armi biologiche. Nessun consigliere propose una minaccia in tal senso per scoraggiare iniziative della Germania o del Giappone.

I legami di Roosevelt con il programma per lo sviluppo di armi biologiche divennero ancora più deboli nel maggio del 1944, quando Stimson e McNutt gli raccomandarono di abolire il WRS e di nominare Merck consigliere di Stimson. Il Presidente approvò subito questa riorganizzazione, decisione che può averlo allontanato ulteriormente dalla iniziativa segreta.

In seguito, in luglio, il capo del gabinetto militare, ammiraglio William D. Leahy e numerosi altri consiglieri intavolarono, alla presenza di Roosevelt, quella che più tardi Leahy definì «un'animata discussione sulle armi batteriologiche». Il dibattito era incentrato sul possibile primo uso di quelle armi per distruggere il raccolto di riso in Giappone. Leahy scrisse più tardi che per lui l'idea era ripugnante; Roosevelt non prese posizione. Il Presidente non dichiarò mai se avrebbe lanciato un attacco biologico per rappresaglia contro un primo uso da parte dell'Asse o se avrebbe autorizzato il primo uso contro il Giappone. (All'epoca circolavano voci che il Giappone avesse impiegato armi biologiche contro la Cina.) In netto contrasto con le sue affermazioni pubbliche che gli Stati Uniti non avrebbero compiuto attacchi con i gas, Roosevelt trasmise così a Truman un'ambigua eredità riguardante le armi biologiche.

Due settimane dopo l'insediamento di Truman alla Casa Bianca, nell'aprile del 1945, e il giorno dopo che al Presidente era stato sottoposto un lungo resoconto sulla bomba atomica, il segretario Stimson ricevette una nota dal suo consigliere speciale Harvey H. Bundy. Questi comunicava che Merck e molti altri membri del programma per la guerra biologica intendevano proporre l'uso di agenti chimici contro i raccolti giapponesi. «È davvero un passo serio», avvertì il consigliere «e forse voi ne vorrete parlare al Presidente». Stimson mandò un messaggio a Marshall chiedendogli di incontrarlo, ma senza urgenza.

Da questo momento fino alla conclusione della guerra, il centro di interesse si spostò dagli agenti batteriologici ai defolianti per i raccolti. Scienziati americani assicurarono che gli agenti chimici non erano velenosi per le persone; il Judge Advocate's Office (l'ufficio competente della magistratura) concluse che il loro uso sarebbe stato legittimo

Un disegno unito a una richiesta di brevetto del 1955 di Robert L. Le Tourneau mostra una «bomba esplosiva per disperdere aerosol tossici e insetticidi». Il progetto deriva forse da quelli di bombe al carbonchio proposti durante la seconda guerra mondiale, quando Le Tourneau lavorava allo sviluppo di armi biologiche. I dettagli della bomba al carbonchio sono tuttora segreti.

I consiglieri del gruppo di ricerca sulle armi biologiche visitano Camp Detrick. George W. Merck, presidente della Merck & Co., Inc., è al centro. Gli altri consiglieri, da sinistra, sono il direttore scientifico Ira L. Baldwin, il capitano Nathaniel S. Prime, il generale di brigata W. A. Borden, il contrammiraglio Julius Zurer, il comandante William B. Sarles, il colonnello Woolpert e il tenente colonnello Norman Pyle.

perché essi non erano tossici e perché gli Stati Uniti, come nazione in guerra, «hanno il diritto di privare il nemico di cibo e di acqua, e di distruggere le fonti di approvvigionamento nei suoi campi».

Stimson, benché profondamente turbato dall'uccisione in massa di civili che i bombardamenti americani avevano già causato, sembrava disposto ad accettare l'avvelenamento dei raccolti in Giappone. Dato che il generale Marshall progettava di impiegare i gas contro le truppe giapponesi, probabilmente anche lui non sarebbe stato contrariato dalla scelta della tattica dell'avvelenamento dei raccolti. In maggio e giugno, un generale dell'aviazione elaborò un piano per distruggere i raccolti di riso cospargendo di tiocianato d'ammonio le aree di produzione vicino a sei delle città più importanti: Tokyo, Yokohama, Osaka, Nagoya, Kyoto e Kobe. Il comandante dell'aviazione, generale Henry H. Arnold, respinse il piano per ragioni tattiche più che morali, ritenendo che bombardare città e industrie giapponesi avrebbe avuto «un impatto più rapido e più sicuro».

In tutte le discussioni successive sull'impiego di defolianti contro il Giappone, i militari misero al primo posto i problemi tattici, proprio come avevano fatto Arnold e i suoi collaboratori. Alcuni si chiesero se le scorte disponibili di agenti chimici sarebbero state sufficienti; altri pensarono che la distruzione del raccolto di riso del 1945 non avrebbe avuto alcun effetto fino al 1946. Si riteneva che per quella data la guerra sarebbe stata vinta e le forze americane di occupazione si sarebbero trovate l'ulteriore fardello di dover nutrire una popolazione civile affamata.

Il 3 agosto, tre giorni prima del lancio della bomba atomica su Hiroshima, il luogotenente generale Ira C. Eaker, vice di Arnold, chiese un rapporto completo sulla distruzione aerea dei raccolti, che descrivesse in dettaglio le possibilità operative dell'aviazione militare, i migliori agenti chimici disponibili e le migliori tecniche per la loro dispersione. Egli ottenne il rapporto il 10 agosto, il giorno successivo al bombardamento di Nagasaki. Quattro giorni dopo, la guerra nel Pacifico era finita.

L'arsenale statunitense di armi batteriologiche, del cui sviluppo l'opinione pubblica nazionale era all'oscuro, non fu dimenticato nei mesi finali della guerra. Già in una fase precedente del programma, uno dei comandanti dell'esercito aveva suggerito che l'Amministrazione potesse prendere in considerazione il primo impiego di queste armi contro il Giappone. Più tardi, le discussioni su eventuali rappresaglie condussero alla conclusione che, nel caso in cui il Giappone avesse violato il Protocollo di Ginevra e fatto ricorso ai gas, gli Stati Uniti avrebbero dovuto essere pronti a rispondere sia con i gas sia con armi batteriologiche. L'ammiraglio Donald B. Duncan, membro dello Stato Maggiore, fece notare che in alcune situazioni l'attacco batteriologico avrebbe potuto essere più efficace di quello con i gas.

In ogni modo, le convinzioni americane sull'accettabilità della guerra biologica non furono mai messe alla prova nella seconda guerra mondiale. La decisione finale sull'impiego delle armi biologiche sarebbe caduta su Truman; egli probabilmente si sarebbe affidato al consiglio del generale Marshall, che ammirava molto, e del segretario Stimson, che considerava un uomo retto. Avendo già approvato l'uso delle bombe atomiche sulle città giapponesi, questi consiglieri forse non avrebbero sollevato obiezioni all'avvelenamento dei raccolti di riso per costringere i giapponesi alla resa.

Sarebbe stato più difficile sostenere la guerra batteriologica, con lo spettro di epidemie e veleni invisibili. Alcuni anni più tardi, comunque, in una lettera a un collaboratore, Truman fece capire che, se la guerra nel Pacifico si fosse protratta fino oltre la metà di agosto, egli avrebbe impiegato sia gli agenti batteriologici sia quelli chimici, e che, in effetti, l'uso di armi atomiche che egli aveva approvato era stato molto peggiore.

Gli Stati Uniti continuarono lo sviluppo e la creazione di riserve di armi biologiche fino al 1969, quando il presidente Nixon, in risposta ai sentimenti pacifisti nati all'epoca della guerra in Vietnam, dichiarò solennemente che avrebbe annullato i programmi e fatto distruggere le riserve. Tre anni più tardi, gli Stati Uniti e più di 100 altre nazioni firmarono la Convenzione sulle armi biologiche e le tossine, e molte di esse si trovarono d'accordo nel bandire completamente la guerra biologica.

All'inizio di questo decennio, gli Stati Uniti cominciarono a sospettare che l'Unione Sovietica avesse un attivo programma di sviluppo di armi biologiche. Come risposta gli Stati Uniti intensificarono il loro programma. I sospetti nacquero nel 1979 da una epidemia di carbonchio che poteva aver avuto origine in un'installazione a Sverdlovsk che gli Stati Uniti asserivano essere un laboratorio militare (ma fra i due fatti potrebbe non esserci in realtà alcuna connessione).

Anche la cosiddetta pioggia gialla, che, secondo vari rapporti, cadde nel Laos e in Kampucea, alimentò sospetti. Il governo degli Stati Uniti sostiene che essa è una micotossina fornita dai sovietici al Vietnam, ma alcuni esperti affermano trattarsi semplicemente di escrementi di api (si veda l'articolo *La «pioggia gialla»* di Thomas D. Seeley, Joan W. Nowicke, Matthew Meselson, Jeanne Guillemin e Pongthep Akratanakul in «Le Scienze» n. 207, novembre 1985).

Le ricerche iniziate clandestinamente durante la seconda guerra mondiale sono perciò cresciute tanto da diventare argomento di dibattito internazionale. Come dovrebbero essere prese le decisioni in tempo di pace? Il passato ci ha insegnato che esse vengono prese a livello politico, non da chi sta in laboratorio. Durante la seconda guerra mondiale gli scienziati fornirono le conoscenze per progettare e sviluppare nuove armi, ma la storia dimostra che essi mancarono di qualsiasi autorità per controllarne lo spiegamento e l'impiego. Entrambi i Presidenti in carica durante la guerra, sebbene investiti dell'autorità formale di comandanti in capo, sapevano assai poco dell'arsenale biologico posto ai loro ordini. Per ironia della sorte, gran parte di quell'arsenale venne sviluppato soprattutto per fronteggiare un paese, la Germania, che non aveva alcuna intenzione di sviluppare capacità offensive nel campo delle armi biologiche.

L'esperienza americana durante la seconda guerra mondiale fa riflettere sul fatto che armi progettate per la deterrenza o la rappresaglia possano sembrare allettanti e moralmente giustificabili anche a scopi offensivi. Una volta che la macchina bellica è entrata in azione, gli scienziati possono non riuscire più a controllare l'uso della tecnologia che loro stessi hanno creato, in particolare nel caso di un conflitto che venga definito una guerra «giusta».

(RI)CREAZIONI AL CALCOLATORE

di A. K. Dewdney

Gli algorompicapo: treni di pensiero seguono binari algoritmici verso la soluzione

Città di Problema e Villasoluzione distano tra loro 160 chilometri. In una bella giornata di giugno, un treno parte a mezzogiorno da ciascuna città e si dirige verso l'altra alla velocità di 80 chilometri all'ora. Contemporaneamente, un calabrone che aveva comodamente sostato sul faro anteriore del treno di Città di Problema si scuote e vola verso Villasoluzione alla velocità di 145 chilometri all'ora. Quando incontra il treno proveniente da Villasoluzione, compie un brusco dietro-front e vola indietro lungo i binari alla stessa velocità, fino a quando incontra il treno proveniente da Città di Problema. Il calabrone continua i suoi avanti e indietro fino a che i treni non si incontrano. Che distanza percorre il calabrone?

La risposta a questo noto rompicapo, che viene data più avanti, è breve; in effetti, è semplicemente un numero. Alcuni rompicapo, che io chiamo algorompicapo, hanno una risposta più complessa. Si tratta di una ricetta, ovvero di un procedimento per determinare un certo stato di cose richiesto. Ci può venir chiesto di dividere il liquido in un contenitore in tre parti uguali attraverso successive operazioni di versamento in contenitori di capacità data. Ci può venir chiesto di far attraversare un fiume a un lupo, una capra e un cavolo attraverso successive operazioni di trasporto. Ci può persino venir chiesto di individuare una moneta falsa (più pesante o più leggera) tra molte buone mediante operazioni consecutive di pesatura su una bilancia.

Un esempio chiarirà la differenza che esiste tra la risposta a un normale rompicapo e quella a un algorompicapo. Per stimolare un corretto percorso mentale, continuerò lungo i binari del primo quesito. Fortunatamente, l'esempio non richiede che i due treni schiaccino il calabrone in un parossismo di vibrazioni infinitesimali. Invece di scontrarsi, i due treni si vedono in tempo per fermarsi. Risuona un lungo fischio, seguito dallo stridore dei freni d'acciaio. (Non so bene perché, mi immagino i due treni come due giganti dell'età del vapore.) Tra i treni si trova un breve tratto di binario e un binario di sorpasso di lunghezza appena sufficiente a ospitare una carrozza o una locomotiva. Ovviamente qualcuno ha commesso un grave errore nel comporre l'orario. Il solo modo che i due macchinisti hanno di trarsi d'impaccio è di utilizzare il binario di sorpasso per consentire ai due treni di oltrepassarsi. In questo problema e in quello che segue sia la locomotiva sia le carrozze hanno un gancio di traino a ciascuna estremità. I due treni hanno anche un frenatore che percorre i binari per agganciare o sganciare determinate vetture.

Asciugandosi il sudore dal viso con un fazzoletto rosso a pallini, il macchinista del treno da Città di Problema esclama: «Non vedo come possiamo superarci a vicenda. Tutto quello che abbiamo a disposizione è quel piccolo binario di sorpasso là!» Il macchinista del treno da Villasoluzione è più ottimista ed espone un piano per permettere ai due treni di superarsi. È possibile? I lettori potrebbero osservare l'illustrazione prima di proseguire nella lettura. A dire il vero, in essa si vedono due treni con cinque vetture ciascuno, ma in un vero algorompicapo i treni hanno sempre n vetture.

Per risolvere questo algorompicapo piuttosto semplice, il lettore avrà ragionato più o meno in questo modo: «Vediamo. Perché non mettere il treno da Città di Problema sul binario di sorpasso, una carrozza per volta? Il treno da Villasoluzione, facendo la spola avanti e indietro, potrebbe tirare ogni volta una nuova carrozza del treno da Città di Problema sul binario di sorpasso in una direzione, superarla di tutta la sua lunghezza e quindi, agganciatala, spingerla fuori del binario di sorpasso per poi ritornare indietro.»

Questa descrizione costituisce il punto di avvio per la soluzione dell'algorompicapo, ma bisogna renderla più chiara e più esplicita; a questo scopo la forma algoritmica è ideale. Sarà utile anche qualche notazione. Ci sono quattro principali sezioni di binario che svolgono un ruolo nella soluzione: il tratto da Città di Problema al binario di sorpasso (A), il tratto tra i due scambi per il binario di sorpasso (B), il binario di sorpasso stesso (B') e il tratto dal binario di sorpasso a Villasoluzione (C). Un'istruzione come «avanti fino a A» significa che il treno da Villasoluzione, composto dalla locomotiva e dalle carrozze al momento agganciate, si muove in avanti finché si trova interamente sul tratto A del binario.

Entrambi i treni hanno n carrozze, naturalmente. La locomotiva del treno da Città di Problema è indicata con $P1$ e le sue vetture con $P2$, $P3$ e così via. L'istruzione «agganciare Pk» significa che il treno da Villasoluzione ha attualmente un'estremità adiacente al k-esimo elemento del treno da Città di Problema; il treno da Villasoluzione avanza lentamente verso Pk e si aggancia con quel delizioso rumore, derivante dalla sua andatura strascicata, che chi ha vissuto vicino alle strade ferrate ricorda bene.

Ora si può dare l'algoritmo:

```
sganciare il treno P
for k = 1 to n + 1
    avanti fino a A
    agganciare Pk
    indietro fino a C
    avanti fino a B'
    sganciare Pk
    indietro fino a C
    avanti fino a A
    indietro fino a B'
    agganciare Pk
    indietro fino a C
    sganciare Pk
agganciare il treno P
```

Dapprima il treno da Città di Problema è completamente sganciato e diviso in unità separate: una locomotiva e n carrozze. Quindi l'algoritmo entra in un ciclo in cui si continuano a ripetere gli stessi 11 passi. Il treno da Villasoluzione compie tutto il lavoro. Senza mai sganciare nessuna delle sue carrozze, il treno avanza verso la sezione A dei binari, dove aggancia la prima ($k = 1$) unità del treno da Città di Problema, la locomotiva. Tira la locomotiva indietro fino alla sezione C e, azionando lo scambio del binario di sorpasso, la spinge fino a B' dove la sgancia. Il treno da Villasoluzione, poi, torna indietro di nuovo fino alla sezione C, aziona lo scambio e avanza verso la sezione A. Successivamente torna indietro nel binario di sorpasso, aggancia di nuovo la locomotiva di Città di Problema e la spinge fuori dal binario di sorpasso fino alla sezione C, dove la sgancia. Il treno, poi, ripete la stessa successione di passi per ciascuna delle vetture del treno da Città di Problema, una alla volta. Quando il ciclo principale è terminato, il treno da Città di Problema si trova dalla parte di Villasoluzione rispetto al treno proveniente da quella stazione. Non appena si sarà ripreso dalla lunga operazione algoritmica, sarà pronto per sbuffare verso la sua meta.

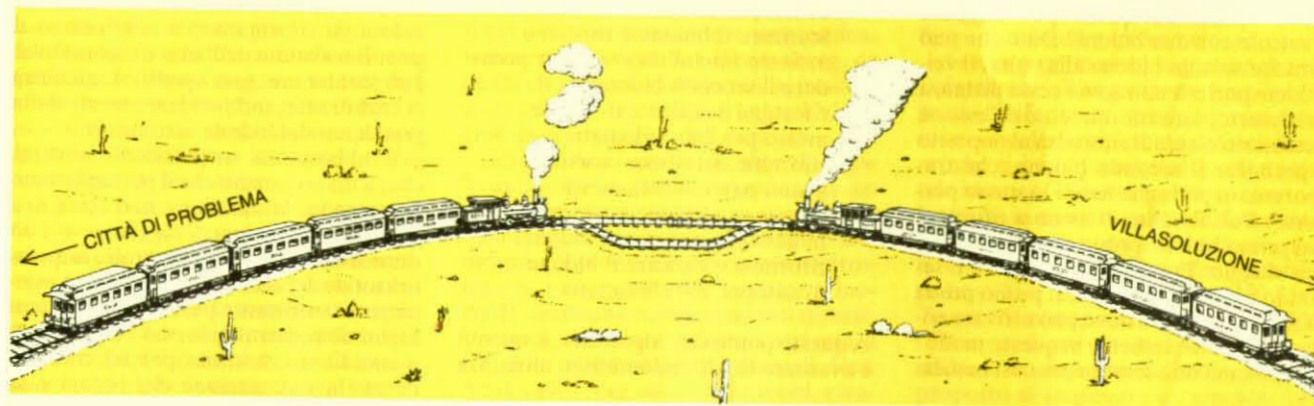
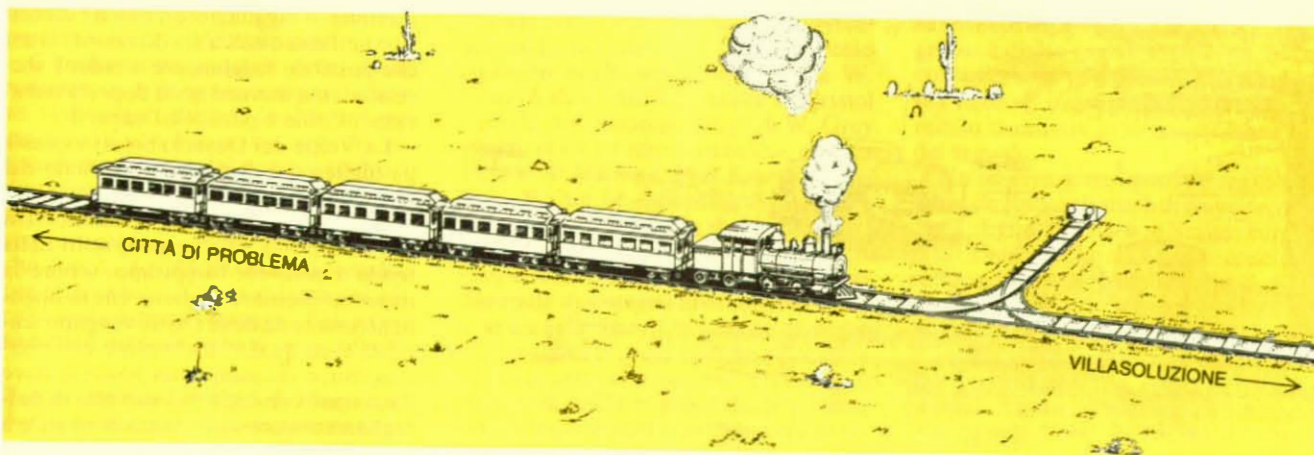
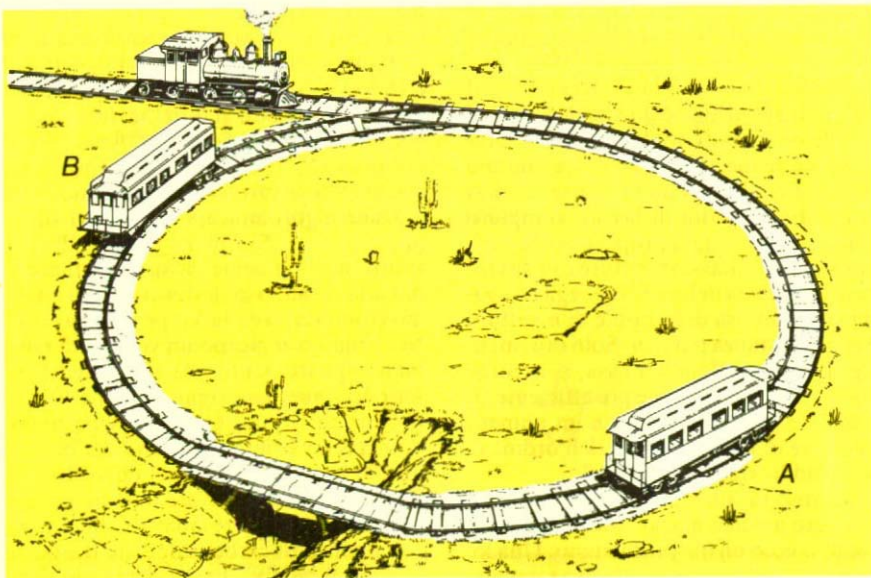
Il successivo esempio di algorompicapo si presenta non appena il treno da Città di Problema completa il suo viaggio per Villasoluzione. Mentre osserva scorrere il panorama, il macchinista si rende conto con orrore di aver dimenticato il p
tornare indietro. Pone mano ai freni e porta il massiccio treno di n carrozze a fermarsi dolcemente. Tornare a marcia indietro fino a Città di Problema non è una bella prospettiva. Fortunatamente il macchinista vede un breve binario morto, con uno scambio per ciascuna direzione di marcia e lungo abbastanza da contenere una carrozza. Ispirato dalla precedente prodezza algoritmica del macchinista di Villasoluzione, il nostro umile eroe non si lascia sopraffare.

Tanto per cominciare, il macchinista di Città di Problema traccia uno schema del
scrive un algoritmo e lo verifica accuratamente, facendo scorrere avanti e indietro il dito sulla carta e borbottando tra sé e sé. Inizia, poi, una laboriosa serie di sganciamenti, deviazioni, riagganciamenti, che lo portano, infine, a far sì che tutto il treno inverta la direzione di marcia. Non solo è stato completamente rovesciato l'ordinamento della locomotiva e dei vagoni, ma la locomotiva stessa e i vagoni sono stati ruotati. Come ha fatto? Pubblicherò le migliori tra le prime soluzioni inviatemi. In questo caso, per migliori si intendono sia le più chiare sia le più intelligenti. Nel problema c'è più di quanto non sembri a prima vista.

Mentre sta avviandosi a recuperare il suo pranzo, il macchinista di Città di Problema si rende conto che la sua soluzione ha richiesto una quantità enorme di energia. Il carbone di scorta è quasi finito. La quantità di lavoro compiuto dal suo treno è stata proporzionale a n^2: misurando la massa in vetture e la distanza in lunghezza delle vetture, egli ha spostato una massa dell'ordine di n vetture per una distanza pari a n. Solo ora, mentre ritorna sbuffando a casa, si accorge che esiste una soluzione più efficiente. È possibile voltare il treno in un numero inferiore di passi, diciamo dell'ordine di $n^{3/2}$ o addirittura di $n \times \log n$?

Al fine di allenare ulteriormente gli intelletti nella nobile arte degli algorompicapo, ecco un bel problemino. Una locomotiva raggiunge un percorso circolare che contiene due vetture vuote. Tra le vetture c'è un ponte, abbastanza robusto per sostenere una carrozza, ma non la locomotiva. Il macchinista deve invertire la posizione delle vetture. Come nel precedente algorompicapo, tutto il lavoro di aggancio e sgancio viene effettuato a treno fermo. (In altre parole, non sono permesse «manovre volanti».) Il ponte non è più lungo di una vettura e la locomotiva deve trovarsi fuori dal percorso circolare al termine del lavoro. La soluzione di questo problema è un semplice algoritmo senza cicli che verrà spiegato ai lettori il mese prossimo.

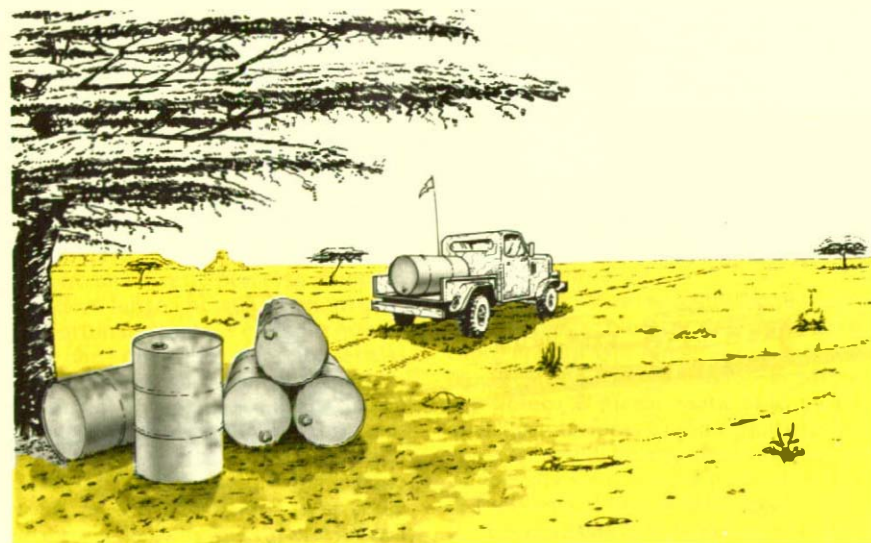
I due algorompicapo successivi sono nuovi per me. Si rimane in tema di trasporti, ma l'indagine passa a una marcia più alta, usando gli autocarri. Io li chiamo i rompicapo della Volpe del Deserto. Nel primo rompicapo un veicolo di pattuglia appartenente alla Volpe del Deserto ha un serbatoio che contiene 40 litri di benzina. Il suo serbatoio viene riempito con il carburante di uno dei bidoni da 200 litri accumulati nel deposito. Se il veicolo può trasportare solo un bidone per volta e percorre quattro chilometri con un litro di carburante, che trasporti o meno il bidone, fino a dove può arrivare prima di restare a secco? La risposta sicuramente dipende dal numero di barili
spirito di generalità degli algorompicapo, il numero di bidoni è indicato con n.

Se n è uguale a 1, la risposta è facile. Il veicolo riempie il serbatoio con il car-

Come fanno i treni a oltrepassarsi?

Come fa il treno a voltarsi per andare a recuperare il pranzo del macchinista a Città di Problema?

La locomotiva deve invertire la posizione delle carrozze senza attraversare il ponte pericolante

burante dell'unico bidone, lo carica e si avvia di pattuglia sotto il caldo sole del deserto. Facendo rifornimento quando è necessario, il veicolo percorrerà ovviamente 800 chilometri prima di restare senza benzina. Per quanto può viaggiare il veicolo con due bidoni? Dato che può caricare solo un bidone alla volta, il veicolo ne porterà uno a una certa distanza nel deserto, farà rifornimento da esso se necessario e quindi ritornerà al deposito a prendere il secondo bidone, che trasporterà in pattuglia lungo lo stesso percorso. Può fermarsi o meno a rifornirsi dal primo bidone prima di continuare il suo viaggio. Può persino scaricare il secondo bidone e prendere il primo prima di ripartire. Fino a dove può arrivare andando avanti e indietro in questo modo? Il problema non sembra più così banale.

Ecco un breve algoritmo che garantisce al veicolo un viaggio di 960 chilometri. L'autocarro parte dal deposito:

rifornirsi e caricare il primo bidone
avanti per 160 chilometri
scaricare il bidone e rifornirsi
indietro fino al deposito per prendere il secondo bidone
rifornirsi e caricare il bidone
avanti per 160 chilometri
rifornirsi dal primo bidone
avanti per 160 chilometri
scaricare il bidone e rifornirsi
indietro fino al primo bidone
rifornirsi e caricare il bidone
avanti per 160 chilometri

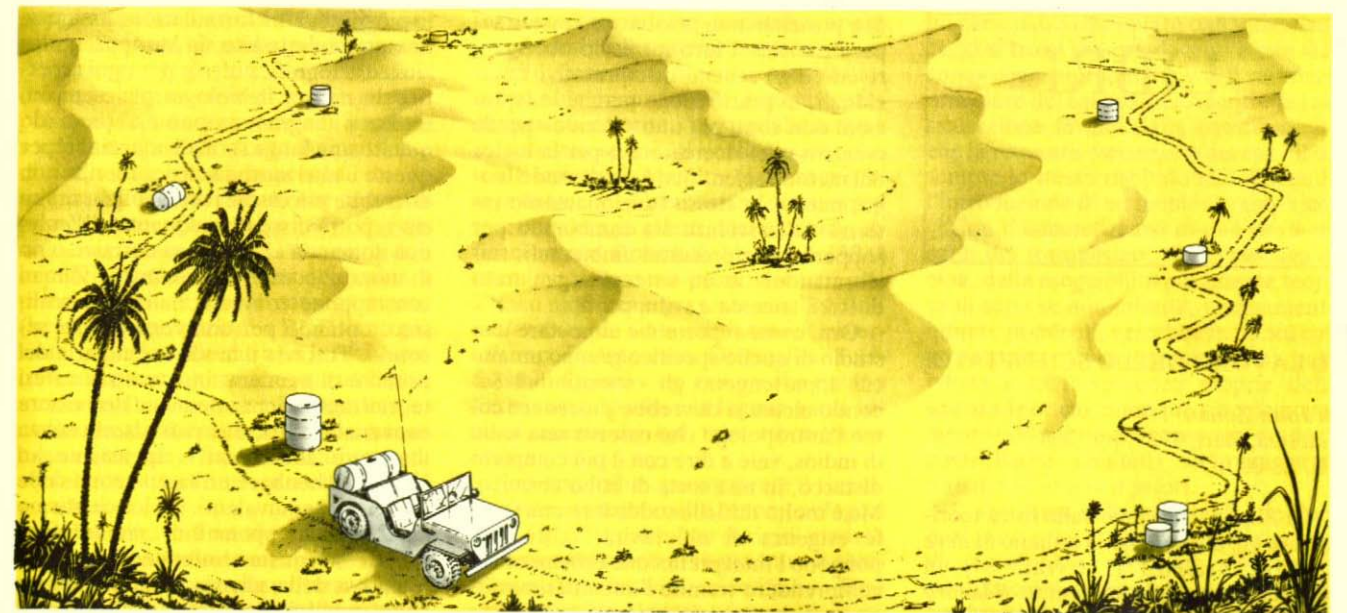
A questo punto dell'algoritmo, il veicolo è avanzato di 320 chilometri in tutto. Ha con sé due bidoni: il primo contiene 40 litri e il secondo ne contiene 120. Il veicolo prosegue lungo le dune dopo che il guidatore ha fatto rifornimento dal primo bidone (lasciandolo vuoto) e ha caricato il secondo. Viaggia, quindi, per altri 640 chilometri prima di restare senza carburante. Poiché, rispetto a quando il barile era uno solo, la distanza percorsa aumenta di soli 160 chilometri, il lettore potrebbe sospettare che esista un algoritmo migliore. In effetti, si può arrivare a quello che il nemico della Volpe del Deserto descriverebbe forse come un «numero bestiale» di chilometri.

Fino a dove può arrivare la Volpe del Deserto se il deposito contiene *n* barili? Nello spirito dell'algorompicapo, non sarà sufficiente una sola risposta numerica o una formula. Si dovrà specificare l'algoritmo che porta alla risposta.

In un'altra parte dell'enorme territorio sabbioso che la Volpe del Deserto chiama (temporaneamente) la sua casa, si trova il percorso di una pattuglia che viaggia in modo molto diverso: un aereo lancia *n* bidoni di carburante in punti arbitrari lungo un percorso circolare chiuso. Ogni bidone contiene una certa quantità di benzina che può variare di molto da un bidone al successivo. Dato che la pattuglia si trova in territorio nemico, l'auto e il pilota vi vengono paracadutati. Essi atterrano vicino a uno dei bidoni, si riforniscono e cominciano il giro. Il serbatoio dell'auto è espandibile: può contenere una qualsiasi quantità di carburante, indipendentemente dalla grandezza del bidone incontrato.

È abbastanza sorprendente notare, che, a mano a mano che il pattugliamento procede, la macchina non resta mai senza benzina prima di raggiungere il bidone successivo. In altre parole, la quantità totale di carburante nei bidoni è esattamente sufficiente perché l'auto di pattuglia completi il giro; né di più né di meno. La cosa è strana, perché, come ho detto, la sistemazione dei bidoni e la quantità di carburante in essi contenuta sono del tutto arbitrari. Per esempio, è perfettamente possibile che l'auto cominci il suo viaggio da un bidone che non contiene abbastanza carburante per permetterle di raggiungere il bidone successivo nell'una o nell'altra direzione. È anche possibile raggiungere il bidone successivo, ma trovarsi poco dopo in imbarazzo. Come è possibile l'impresa?

La Volpe del Deserto osserva queste pattuglie ed è decisamente turbato dal loro comportamento. Il nemico, notoriamente pazzo, è anche fortunato? Un suo sottoposto, che passa le notti nella tenda a risolvere rompicapo, scopre la risposta: «Sembra, signore, che indipendentemente da dove i barili vengano scaricati e da quanto carburante essi contengano, ci sia sempre un posto da dove l'auto può cominciare il suo giro di pattugliamento con la certezza di riuscire a completare il viaggio».

Lascio ai lettori il compito di scoprire

Quanta strada può fare il veicolo di pattuglia con n barili di carburante?

Da dove deve partire l'auto di pattuglia per essere sicura di poter completare il circuito?

il segreto. Ovviamente ha a che fare con un algoritmo. Bisogna dapprima trovare una procedura passo per passo che disponga almeno un barile dal quale l'auto possa partire per cominciare (e completare) il suo giro di pattuglia. L'algoritmo deve anche specificare la direzione nella quale l'auto deve viaggiare: andare nel senso sbagliato potrebbe essere disastroso, come nel caso degli algorompicapo dei treni. Pubblicherò le migliori tra le prime risposte inviatemi.

All'inizio dell'articolo ho promesso di rivelare la risposta al rompicapo del calabrone e del treno. Chi propone rompicapo è spesso pronto a fuorviare i solutori, specie quando la risposta è semplice. Descrivendo il numero infinito di dietro-front del calabrone, ho deliberatamente cercato di confondere i lettori. Si pensi soltanto che il calabrone è in volo per tutto il tempo necessario ai due treni per incontrarsi. Essi percorrono metà della distanza tra Città di Problema e Villasoluzione in un'ora esatta. Il calabrone, volando a 145 chilometri all'ora, percorre, quindi, 145 chilometri.

C'è una storia apocrifa circa questo rompicapo, sulla quale non posso giurare. Si dice che qualcuno lo abbia proposto a John von Neumann, uno dei più grandi matematici del nostro secolo. Si dice che egli abbia risposto subito: «145 chilometri, ovviamente». «Ah, - disse il suo torturatore, - avrei dovuto saperlo che avresti trovato un sistema semplice per risolverlo!» «Quale sistema semplice?» replicò von Neumann. Egli aveva sommato mentalmente la serie infinita.

Le reazioni all'articolo di maggio sui veicoli di Valentino Braitenberg sono state entusiastiche e sono andate da suggerimenti pratici a riflessioni filosofiche. I veicoli sono astratti: una piastra sostiene un circuito stampato, una batteria di sensori (generalmente due occhi) frontali e due motori che comandano le ruote posteriori. Sono sufficienti i circuiti più semplici per rendere sorprendentemente vari i veicoli di Braitenberg, da quelli che fissano globi luminosi in silenziosa adorazione, a veicoli aggressivi che emergono rombanti dall'oscurità.

Joseph A. Coppola di Sherrill, New York, sostiene che si può costruire un vero veicolo con 30 dollari, usando per il montaggio una piastra di compensato, motori già pronti e componenti elettronici. Si può ricablare il veicolo in modo da verificare comportamenti diversi. I lettori interessati possono scrivere a Coppola, 304 Park Street, Sherrill, New York 13461, per ottenere progetti, elenco dei componenti e fornitori. È ragionevole immaginare che di qui a un anno ci siano veicoli funzionanti che possano essere mandati nella metropoli a divertirsi sotto i lampioni?

Tom Napier di Dresher, Pennsylvania, mi ha ricordato un raffinato veicolo costruito negli anni cinquanta da W. Grey Walter, dell'Università di Bristol (si veda *An Imitation of Life* di W. Grey Walter in «Scientific American», maggio 1950 e *A Machine That Learns* di W. Grey Walter in «Scientific American», agosto 1951). Il veicolo si chiamava tartaruga. Secondo quanto dice Napier, «aveva una sola ruota motrice e usava due ruote di sostegno liberamente rotanti su un asse fisso. Un motore di sterzo faceva girare la ruota motrice e il motore di guida attorno a un asse verticale. Il motore di guida e quello di sterzo giravano continuamente ma erano regolati a mezza velocità o a tutta velocità da due relè, controllati da due valvole a vuoto alimentate da una fotocellula che ruotava con la direzione di sterzo. Dato che il motore di guida e quello di sterzo avevano due stati e non si spegnevano mai, la tartaruga si muoveva verso le fonti di luce secondo un percorso cicloidale.»

Se non è prematuro suggerire che chi si interessa all'argomento funzioni da punto di riferimento per le comunicazioni sui veicoli, farei il nome di Gary Blauer, che si è spontaneamente offerto per tale incarico. Blauer sarebbe contento di mettersi in contatto con chi fosse interessato a leggere (e a scrivere) sui veicoli. Gli si possono inviare comunicazioni presso Dain Bosworth, Inc., 100 Dain Tower, Minneapolis, Minnesota 55402.

Per produrre un comportamento oscillante in un veicolo di Braitenberg, ho proposto di impiegare un circuito fatto di quattro neurodi: semplici elementi formali di calcolo che sono l'equivalente veicolare dei neuroni. Frank Palmer, di Chicago, precisa che bastano due neurodi invece di quattro. Nel suo schema il cavo elettrico che porta il segnale di ingresso è diviso in due e ciascuna metà è collegata a un neurode di soglia 1. Uno dei neurodi inibisce l'altro. Il neurode inibito trasmette in uscita ciò che resta del segnale.

Un'osservazione curiosa e raggelante viene da Robert Baruch di Bayside, New York. Egli concorda con Braitenberg sul fatto che osservatori esterni concluderanno che anche semplici veicoli manifestano emozioni. Baruch ipotizza che forse la percezione delle nostre emozioni è analoga. Scrive Baruch: «È possibile che il cervello umano, con la miriade dei suoi meccanismi di retroazione, consideri se stesso dotato di emozioni e di un concetto chiamato "mente", anche se essi non sono altro che circuiti.»